

STRUKTUR UND EXPRESSION
SPEZIFISCH REGULIERTER GENE
IN BINDEGEWEBEN DES HUHNS

Ernst Pöschl

*Ass.
München
1984*



Digitized by the Internet Archive
in 2026

awi-druck
Dissertations-Druck
Müllerstraße 22 München 5
Telefon 260 40 05

STRUKTUR UND EXPRESSION
SPEZIFISCH REGULIERTER GENE
IN BINDEGEWEBEN DES HUHN

Dissertation
der Fakultät für Biologie
der Ludwig-Maximilians Universität.
München

Vorgelegt von
Ernst Pöschl
aus München

München 1984

1. Berichterstatter : PD Dr. K. Moritz

2. Berichterstatter : Prof. R. Schweyen

Sondervotum : Prof. K. Kühn

Tag der mündlichen Prüfung: 6. 11. 1984

Meiner Familie

Meinen Eltern und Großeltern

INHALTSÜBERSICHT

A. EINLEITUNG	1
B. AUFGABENSTELLUNG	17
C. ERGEBNISSE	18
1. Charakterisierung des cDNA Klones p550	18
1.1 Klonierung von Fragmenten des cDNA Klons p550	18
1.2 Sequenzierung der Klone 550/112-1 und -8	24
1.3 Orientierung und Proteinsequenz des cDNA Klones p550	27
1.4 Gewebespezifität des Klons p550	33
1.5 Direkte Identifizierung des p550 Proteins durch Hybrid Selektion	36
2. Konstruktion gewebespezifischer Genbanken	43
2.1 Prinzip der Konstruktion von Genbanken	43
2.2 Isolation genomischer DNA und Partialspaltung mit EcoRI	46
2.3 Klonierung in λ Ch 4A	49
2.4 In Vitro Packaging der Genbanken	50
2.5 Background der Genbanken	54
3. Isolierung und Restriktionskartierung genomischer Klone des p550 Gens	55
3.1 Prinzip der Isolierung genomischer Klone	55
3.2 Screening der 'Sterna' Genbank mit dem cDNA Klon p550	56
3.3 Restriktionskartierung der Klone λ St 550-1 bis -5	59

3.4	Restriktionskartierung der Klone λ St 550-3 und -5	61
3.5	Feinkartierung der Eco RI-Spaltstellen	67
3.6	Vergleich der Restriktionskarten von Genom und klonierter Region	68
3.7	Methylierungsmuster des klonierten Genbereichs	71
4.	Subklonierung und Isolierung weiterer genomischer Klone	74
4.1	Subklonierung verschiedener Genbereiche	74
4.2	Eingrenzung des kodierenden Bereichs	77
4.3	Versuche zur Ausweitung des bisher isolierten Genbereichs	80
4.4	Isolierung des p550-Gens aus der DNA dedifferenzierter Chondrocyten	82
5.	Feinanalyse des p550 kodierenden Genbereichs	83
5.1	Sequenzierung des p550 kodierenden Genbereichs	83
5.2	Exon-Intron Struktur des p550 kodierenden Bereichs	87
6.	Charakterisierung des 5'-terminalen Bereichs des p550 Gens	91
6.1	Restriktionskartierung des Klons pSt 608	91
6.2	Grobe Lokalisierung der kodierenden Bereiche	94
6.3	Sequenzierung von Teilen des Klons pSt 608	94
6.4	Analyse der Sequenz von pSt 608	105
7.	Untersuchung der cDNA Klone p2200 und p1300	112
7.1	Konstruktion der cDNA Klone p2200 und p1300	112

7.2	Restriktionskartierung von p2200 und Subklonierung von Fragmenten	112
7.3	Sequenzierung des Klons p2200	114
7.4	Genomorganisation des p2200 kodierenden Bereichs	117
7.5	Orientierung und Expression des p2200 Gens	120
7.6	Restriktionskartierung und Sequenzierung des cDNA Klons p1300	122
7.7	Kartierung der p1300 kodierenden genomischen Region	126
7.8	Expression der p1300 spezifischen mRNA	127
D.	DISKUSSION	129
1.	Identifizierung des p550-Genprodukts	129
2.	Isolation des p550-Gens	137
3.	Exonstruktur des p550-Gens	140
4.	Hypothetische Promoterregion des p550-Gens	142
5.	Transkription des p550-Gens	146
6.	Untersuchung weiterer cDNA-Klone	149
7.	Verwendung spezifischer cDNA-Klone	151
E.	ZUSAMMENFASSUNG	152
F.	ABKÜRZUNGEN	154
G.	MATERIALIEN UND METHODEN	155
1.	Materialien	155
2.	Methoden	157
2.1	Zellkultur	157
2.2	DNA Isolierung	157
2.3	Restriktionsspaltungen. Gele, Southernblotting	161
2.4	RNA Isolierung und Northernblotting	163

2.5 Radioaktive Markierung von DNA/RNA und Hybridisierung	164
2.6 Konstruktion der Genbanken	167
2.7 Screening der Genbanken	168
2.8 Klonierung in Plasmiden	172
2.9 Klonierung in M13 und Charakterisierung von M13 Klonen	173
2.10 Sequenzierungsmethoden	174
2.11 Hybrid Selektion und In Vitro Translation	177
2.12 Computerprogramme und Datenbanken	178
H. LITERATURANGABEN	180

A. EINLEITUNG

Die Entwicklung hochspezialisierter Zellen aus undifferenzierten pluripotenten Mesodermzellen ist entscheidend für die Ausbildung funktionell unterschiedlicher Gewebe und Organe in einem Embryo. Neben organspezifischen Zelltypen entwickeln sich dabei auch die Zellen der verschiedenen Formen des Bindegewebes, wie Osteoblasten, Chondrocyten und Fibroblasten. Diese Zellen synthetisieren große Mengen extrazellulärer Matrix, deren Zusammensetzung und Struktur verantwortlich ist für die mechanische Stabilität, Elastizität und Permeabilität der einzelnen Gewebe (1).

Verschiedene Faktoren spielen bei der Induktion und der weiteren Steuerung der Differenzierung und später bei der Aufrechterhaltung des differenzierten Zustands eine Rolle. Die genetische Programmierung leitet diesen Prozeß ein, und durch Zell-Zell Wechselwirkungen wird Zellmorphologie und weitere Differenzierung gesteuert (2). Bei der folgenden Ausbildung der Gewebe und der Aufrechterhaltung des differenzierten Zustands spielen im wesentlichen Zell-Matrix Wechselwirkungen eine Rolle, die neben der Funktion bei der Cytodifferenzierung einen wesentlichen Einfluß auf Zellwanderung und Zellwachstum haben (3). Dieses Durchlaufen verschiedener Stadien läßt sich bei den Zellen des Bindegewebes an Veränderungen in Menge und Zusammensetzung der produzierten extrazellulären Matrix besonders leicht verfolgen.

Der Knorpel bietet ein solches, leicht untersuchbares System zur Untersuchung der Frage nach Regulation und Struktur von Matrixkomponenten. Ein Zelltyp, der Chondrocyt, produziert hier in großer Menge eine, in dieser Zusammensetzung, nur im Knorpel zu findende hochspezialisierte Matrix (4). Knorpel bildet während der frühen Embryonalentwicklung das Gerüst für die spätere Skelettbildung und wird, mit Ausnahme kleiner knorpeliger Wachstumszonen, später kalzifiziert und zum Knochen umgebaut. Im entwickel-

ten Organismus findet sich Knorpel in den Gelenken als Druckpolster und in Bereichen des Körpers, in denen hohe Flexibilität und Elastizität verlangt werden. Es lassen sich dabei 3 Formen unterscheiden : der semitransparente hyaline Knorpel in Gelenken, Nasensepten, den Trachealringen und dem Sternum, der fibrilläre Knorpel in Menisken und den Zwischenwirbelscheiben und der elastische Knorpel im Außenohr. Allen gemeinsam ist die große Menge an extrazellulärer Matrix, die fehlende Innervierung und eine geringe Versorgung durch Blutgefäße.

Die Knorpelmatrix ist aufgebaut aus drei Arten von Komponenten. Kollagene und Proteoglykane bilden die mengenmäßig vorherrschenden Gruppen, die im wesentlichen für die Elastizität und mechanische Stabilität des Gewebes verantwortlich sind. Daneben findet sich noch eine Anzahl von Proteinen, die keiner der beiden Gruppen zugehören, und denen spezielle Funktionen im Knorpel zugeordnet werden können. Während die beiden Hauptkomponenten relativ gut beschrieben sind, ist über die vielen minoren Bestandteile wenig bekannt.

Kollagene

Die Kollagene, eine Gruppe genetisch verwandter Glykoproteine, stellen im Gesamtorganismus die mengenmäßig vorherrschenden Proteine dar. Gekennzeichnet sind alle Mitglieder dieser Gruppe durch die Ausbildung eines besonderen Helixtyps - der Tripelhelix - bei der Assoziation von 3 Peptidketten, den α -Ketten (5). Voraussetzung für diese Struktur ist die dauernde Wiederholung der Sequenz Gly-X-Y innerhalb der tripelhelikalen Sequenzen. Dieser spezielle Helixtyp findet sich, außer bei den Kollagenen, noch in Teilen der Acetylcholinesterase und der C1q Komponente des Komplementsystems (6). Ein besonderes Charakteristikum der interstitiellen Kollagene ist die Ausbildung hochmolekularer Aggregate durch eine geordnete Zusammenlagerung der Einzelmoleküle zu Fibrillen und deren Stabili-

Reste werden hydroxyliert und kurze Oligosaccharide (Gal und Glu) werden an Hydroxylysin angeheftet. Beim Transport aus der Zelle werden die terminalen, globulären Prokollagen-peptide abgespalten und damit die Ausbildung der fibrillaren Struktur ermöglicht (7).

Im Knorpel findet sich als Hauptkomponente Kollagen Typ II, neben einer geringen Menge (10 - 20%) minorer, kollagener Proteine (10). Typ II Kollagen wurde zuerst im Sternum von Hühnern nachgewiesen (11), später auch in anderen knorpeligen Geweben und im Glaskörper des Auges. Gemeinsam mit Typ I Kollagen ist es nur im Nucleus pulposus der Zwischenwirbelscheiben zu finden. Kollagen Typ II ist aufgebaut aus 3 identischen, 1020 Aminosäuren langen $\alpha 1(II)$ Ketten. Von den anderen interstitiellen Kollagenen unterscheidet sich dieser Typ hauptsächlich durch die starke Hydroxylierung der Lysin-Reste (etwa 70% aller Lysin-Reste) und einer damit verbundenen starken Glykosylierung (etwa 3% des Gesamtgewichts) Diese Seitenketten stören die Ausbildung großer Aggregate, weshalb im Knorpel deutlich dünnere Fibrillen (10 -20 nm) zu finden sind als beispielsweise bei Typ I-Kollagenfibrillen der Haut (100-200 nm). Die Proteinsequenz von Kollagen Typ II (Rind) ist bisher zu etwa 50% bekannt. Ambiguitäten in der Sequenz zeigen, daß es wahrscheinlich mindestens 2 sehr eng verwandte $\alpha 1(II)$ Peptide gibt (12).

Die Typ II-Fibrillen sind im Vergleich zu Kollagenfibrillen der Haut (Typ I) nicht in dicken Bündeln angeordnet, sondern bilden ein loses Netzwerk im Knorpel. Ihr Beitrag zur Funktion liegt deshalb weniger in der Ableitung großer Zugspannungen, sondern im elastischen Zusammenhalt der Gesamtstruktur.

Neben dem Haupttyp des Typ II Kollagens wurden eine Reihe weiterer minorer Kollagene in verschiedenen Knorpeln nachgewiesen (10). Drei zusätzliche kollagene Peptide, 1α , 2α und 3α , in der Größe der α -Ketten wurden zuerst in humanen, hyalinen Knorpeln gefunden (13), später dann auch im

Sternum von Huhn und in anderen Organismen. Die 1α und 2α Ketten werden der Familie der Typ V Kollagene zugerechnet, während die 3α -Kette Ähnlichkeiten zur $\alpha_1(\text{II})$ Kette aufweist. Bei verschiedenen Organismen und Knorpeltypen finden sich daneben noch zusätzliche disulfid-quervernetzte Kollagentypen, so das 'pM-Kollagen (Typ IX)' im Sternum von Huhn (14), das als die Vorläuferform zweier früher beschriebener kollagener Proteine (HMW und LMW (15)) anzusehen ist. Diese minoren Kollagene scheinen innerhalb der Matrix in einzelnen Domänen und Bereichen vorzuliegen. Bei unterschiedlichen Kulturbedingungen synthetisieren die Chondrocyten zusätzlich weitere kollagene Peptide (16,17), über deren Funktion, ebenso wie für die der anderen minoren Komponenten, noch nichts bekannt ist.

Proteoglykane

Etwa ein Drittel der Trockenmasse des hyalinen Knorpels besteht aus Proteoglykanmolekülen, deren typischer Aufbau in Abb. 2 gezeigt ist (18). An ein Proteincore ist eine große Anzahl von Glykosaminoglykanen und kurzen Oligosaccharidseitenketten gebunden. Die elektrostatische Abstoßung einer großen Zahl negativer Ladungen von Sulfat- und Carboxylgruppen in diesen Seitenketten erzwingt eine starke Auf-faltung und damit einen großen Platzbedarf der Moleküle im Gewebe. Da die Moleküle in der Matrix in einer dicht gepackten, komprimierten Form vorliegen, setzen sie jeder weiteren Kompression einen stetig ansteigenden Widerstand durch die dabei ebenfalls ansteigende Abstoßung der einzelnen Ladungen entgegen. Auf diesem Effekt beruht im wesentlichen die Elastizität und besonders die Druckstabilität des Knorpels.

Die Proteoglykane liegen meist nicht isoliert als Einzelmoleküle vor, sondern bilden zusammen mit Hyaluronsäure große Aggregate. Etwa 50 Proteoglykane sind über eine spezielle Bindungsregion an ein Hyaluronsäuremolekül gebunden. Diese Bindung wird noch weiter stabilisiert durch die 'Link-Proteine'.

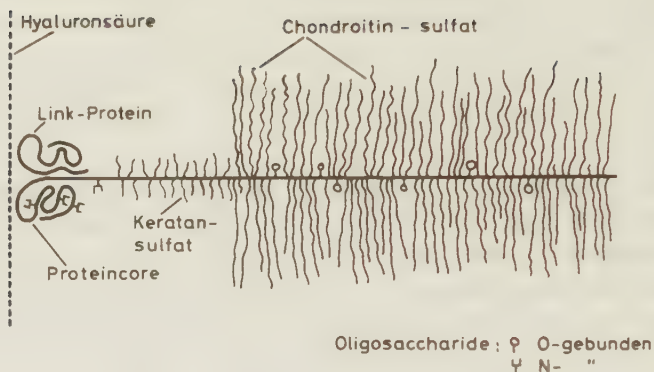


Abb. 2 : Schematischer Aufbau eines typischen Proteoglykanmoleküls des Knorpels (PG-H)

Die meisten Untersuchungen beziehen sich auf den Haupttyp der Proteoglykane des Knorpels mit einem Molekulargewicht von etwa 2.5×10^6 d. Der Anteil des Coreproteins beträgt nur etwa 8% der Gesamtmasse, und dessen Molekulargewicht wird meist mit etwa 200 kd angegeben (19). Da für das in vitro translatierte Vorläuferprotein eine Länge von etwa 340 kd bestimmt wurde, scheint das Protein in einer Reihe von posttranslationalen Modifikationsschritten verändert zu werden (20). Über die Struktur des Coreproteins ist, bis auf wenige fragmentarische Sequenzdaten (21), noch sehr wenig bekannt. Ein typisches Molekül enthält etwa 100 Chondroitinsulfat- und 15 - 30 Keratansulfatseitenketten, die in Gruppen auf dem Molekül angeordnet sind, wie es in Abb. 2 gezeigt ist. Verteilt auf dem Molekül finden sich noch N- und O- glykosidisch gebundene Oligosaccharide.

Dieser Grundaufbau unterliegt verschiedenen Variationen. So lassen sich sowohl unterschiedliche Längen des Proteincores nachweisen als auch Unterschiede in der Zusammensetzung und dem Verteilungsmuster der Glycosaminoglykane. Mindestens drei Untergruppen sind dabei unterscheidbar. Neben dem Haupttyp der Proteoglykane (PG-H) scheint einem Teil

der Moleküle die Hyaluronsäurebindungsstelle zu fehlen, während sich eine andere Subpopulation vom Prototyp durch das Fehlen der Keratansulfatseitenketten unterscheidet. Möglicherweise sind die Proteincorcs dieser Gruppen verwandte, aber individuelle Mitglieder einer Multigenfamilie (22).

Nicht nur die Anzahl und Anordnung der Glycosaminoglykane ist variabel, auch die Länge der Seitenketten unterliegt deutlichen Veränderungen. Während in frühen Phasen lange Chondroitinsulfate und kurze Keratansulfatseitenketten vorherrschen, kehrt sich dieses Verhältnis im Verlauf der Differenzierung um (23).

Neben dem beschriebenen Haupttyp von Proteoglykanen (PG-H) des Knorpels läßt sich durch Gradientenzentrifugation noch eine weitere Gruppe von weniger dichten Proteoglykanen abtrennen (24). Genauer beschrieben wurden dabei das PG-Lt (25) und PG-Lb (26), die aus dem Epiphysenknorpel embryonaler Hühner isoliert wurden. Die Coreproteine (PG-Lt = 100 + 120 kd; PG-Lb = 43 kd) sind im Vergleich mit PG-H deutlich kleiner und die Glykosaminoglykanseitenketten unterscheiden sich sowohl in ihrer Anzahl und Länge, als auch in ihrer Zusammensetzung. Eine Besonderheit des PG-Lt liegt darin, daher eine Assoziation aus zwei unterschiedlichen Peptidketten vorliegt, die durch bakterielle Kollagenase gespalten werden können, damit zumindest einige kollagenähnliche Sequenzen aufweist. Dieser Proteoglykantyp findet sich nicht nur in chondrogenen, sondern auch in nicht-chondrogenen Zellen der Epiphysen, während PG-H spezifisch für den Knorpelzustand ist.

Sonstige nichtkollagene Matrixkomponenten

Zusätzlich zu den Proteoglykanen und Kollagenen ließ sich eine Anzahl weiterer Proteine mit spezialisierten Funktionen innerhalb der Matrix nachweisen. Ein Beispiel sind die 'Link-Proteine' (27), die für die Stabilität der Proteoglykan - Hyaluronsäure Wechselwirkung von entscheidender Bedeu-

tung sind, da sie sowohl an Proteoglykane als auch an Hyaluronsäure binden. Ein Bereich von etwa 25 Monosaccharideinheiten der Hyaluronsäure ist an der Bindung des Link-Proteins und eines Proteoglykanmoleküls beteiligt (Abb. 2). Zwei strukturell unterschiedliche Linkproteine mit Molekulargewichten von 46 kd und 40.5 kd, konnten bisher nachgewiesen werden (28). Diese einmalige Form der Aggregation von Proteoglykanen, Linkproteinen und Hyaluronsäure findet sich nur im Knorpel und scheint von grundsätzlicher Bedeutung für die innere Organisation und Funktion des Knorpels zu sein.

Ein chondrocytenspezifischer Zellerkennungsfaktor, das Chondronektin, findet sich auf der Oberfläche der Zellen. Dieses Protein vermittelt die Anheftung von Chondrocyten an Typ II Kollagensubstrate und unterscheidet sich im Wirkungsmechanismus deutlich von dem des funktionell analogen Fibronektins. Aufgebaut ist dieses 180 kd Glykoprotein aus Untereinheiten von 80 kd (29).

Ein besonders interessantes Beispiel eines minoren Matrixproteins ist ein 35 kd großes Glykoprotein, das Chondrocalcin (30), das aus fötalen Epiphysen isoliert werden konnte. Es scheint, daß dieser Komponente eine wesentliche Funktion bei der Mineralisierung des Knochens zukommt.

Diese Gruppe von Proteinen, die in geringerer Menge in der Matrix des Knorpels auftreten, ist zwar noch wenig untersucht, hat aber wahrscheinlich wichtige Funktionen bei der Zellerkennung, Differenzierung und bei der Stabilisierung der inneren Struktur der Matrix.

Aufbau der Matrix

Eingebettet in diese hochstrukturierte komplexe Matrix liegen die Chondrocyten einzeln oder in Gruppen vor. In früheren Arbeiten wurde angenommen, daß die Chondrocyten in einer Art Tasche weniger dichten Materials auftreten. Aus neueren Untersuchungen läßt sich jedoch schließen, daß die Chondrocyten annähernd gleichmäßig verteilt in die Matrix

eingebettet sind (31).

Die Kollagenfibrillen umgeben als ein Netzwerk die amorphe Masse der Proteoglykan - Hyaluronsäure Aggregate und bieten dadurch Stabilität gegenüber Scher- und Zugbelastungen, die während der Kompression des Gewebes im Gelenkknorpel auftreten. Die Proteoglykanaggregate gewährleisten die Druckstabilität des Gewebes durch den Widerstand, den sie einer Kompression entgegensetzen. Beide Gruppen von Molekülen scheinen über eine große Anzahl von Kontaktstellen nicht nur untereinander, sondern auch mit der Zellmembran der Chondrocyten verknüpft zu sein (32). Diese Kombination von Komponenten mit unterschiedlichen Eigenschaften und deren Organisation innerhalb einer kompakten Matrix stellen eine optimale Anpassung an die funktionellen Bedürfnisse des Knorpels dar.

Neusynthese und Abbau der Matrixkomponenten

Während der Entwicklung und des Wachstums unterliegt die Zusammensetzung der Matrix einem dauernden Umbauprozess, was durch Degradation und Neusynthese der Komponenten erreicht wird. Ein solcher Mechanismus bietet die Möglichkeit einer graduellen Veränderung der Zusammensetzung. Während die Synthese von Proteoglykanen (33) und Kollagenen (34) bereits in vielen Einzelheiten untersucht wurde, ist über die Mechanismen des Abbaus noch wenig bekannt. Eine Reihe nicht genauer charakterisierter proteolytischer Enzyme ist wahrscheinlich verantwortlich für den Abbau der Proteinanteile. Die Kollagene werden von verschiedenen spezifischen und unspezifischen Proteasen abgebaut und unterliegen, zumindest im Gelenkknorpel, einem extrem langsamen Turnover (35).

Regulation der Synthese von Matrixkomponenten

Die Regulation der Synthese von Einzelkomponenten im Verlauf der Differenzierung des Knorpels aus undifferenzierten Mesenchymzellen läßt sich an der Entwicklung der

Flügelanlagen des Huhns gut verfolgen. Zwei getrennte Stufen sind unterscheidbar (Abb. 3): in der morphogenetischen Phase (2-4 Tage) entwickelt sich die noch undifferenzierte Anlage: in der Phase der Cytodifferenzierung (4-5 Tagen) erfolgt dann die Ausbildung des speziellen Phänotyps, sichtbar am Auftreten von Kollagen Typ II und knorpelspezifischer Proteoglykane (36). Diese Komponenten werden nun in großer Menge synthetisiert und als Matrix um die Zellen abgelagert. Die Synthese von Kollagen Typ I der mesenchymalen Vorläuferzellen wird dabei abgeschaltet.

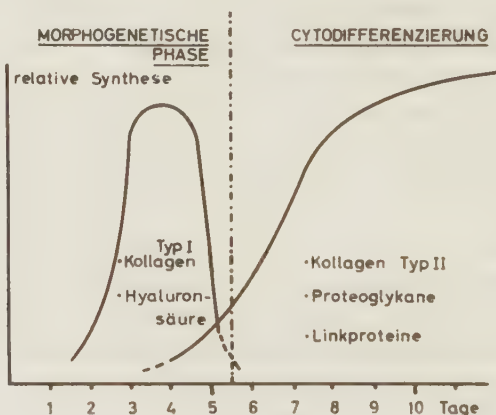


Abb. 3 : Entwicklung der 'Limb buds' (36)

Bereits in der ersten Phase der Entwicklung der noch weitgehend undifferenzierten Zellen lassen sich Gruppen von Zellen unterscheiden, deren Differenzierung zu Muskel- oder Knorpelzellen bereits determiniert ist. Diese Positionsunterschiede spiegeln die prinzipielle Bedeutung von Zell-Zell Wechselwirkungen in diesem Stadium der Gewebeentwicklung wider. Dieser Einfluß ist auch bei der in vitro Kultur von Mesenchymzellen der morphogenetischen Phase erkennbar: extrem hohe Zelldichten führen zur Differenzierung in Chondrocyten und der Synthese von Knorpelmatrix. Im

Gegensatz dazu, entstehen bei zu geringen Zelldichten jedoch Muskelzellen (36).

Die Chondrogenese läßt sich durch eine Reihe von Faktoren beeinflussen, die dadurch einen Einblick in die Mechanismen der Regulation geben könnten (36). Als erster Hemmstoff der Chondrogenese wurde das Nicotinamid beschrieben, das zu erhöhten intrazellulären Konzentrationen von NAD führt und dadurch anscheinend die Expression der spezifischen Gene blockiert. Vermittelt werden könnte diese Wirkung durch die Synthese von Poly-ADP-Ribose im Zellkern. Dafür spricht auch der Nachweis einer erhöhten Aktivität der beteiligten Enzyme in differenzierenden Chondrocyten.

Auch durch Gaben von Bromdeoxyuridin (BrdU) läßt sich die Chondrogenese hemmen, möglicherweise ebenfalls durch einen Einfluß auf die Synthese von Poly-ADP-Ribose. Da BrdU in die DNA eingebaut werden kann, besonders in repetitive Sequenzbereiche, bietet sich für dessen Wirkung auch ein direkter Einfluß auf regulatorische Elemente spezifischer Gengruppen an (37).

Stimulieren läßt sich die Chondrogenese dagegen durch einen erhöhten cAMP Gehalt (38), der in vivo vielleicht die Folge der massiven Zell-Zell Kontakte ist.

Neben diesen Steuerungsvorgängen, die im wesentlichen auf Zell-Zell Wechselwirkungen beruhen, läßt sich auch ein Einfluß von Zell-Matrix Wechselwirkungen auf die Ausprägung des differenzierten Phänotyps zeigen. Die Zugabe von exogener Hyaluronsäure, die in sehr frühen Stufen der Embryonalentwicklung in großen Mengen gebildet wird, während der Chondrogenese jedoch teilweise abgebaut wird, hemmt die Synthese von Proteoglykanen. Zugabe von Chondroitinsulfat dagegen regt die Synthese der Proteoglykane an (39).

Während in den mesenchymalen Vorläuferzellen nur Kollagen Typ I produziert wird, findet sich im differenzierten Gewebe nur noch Kollagen Typ II. Wird der Knorpel später durch Knochen ersetzt, schalten die Zellen wieder auf die Synthese von Kollagen Typ I um (40). Ein solcher Umschaltprozeß der Synthese verschiedener Kollagentypen läßt

sich besonders gut bei der Kultur von Chondrocyten beobachten (41). Werden Chondrocyten als Monolayer kultiviert, schalten die Zellen von der Synthese des Typ II Kollagens auf die Synthese von Typ I um (Abb. 4). Durch die Zugabe von BrdU läßt sich dieser Vorgang noch erheblich beschleunigen. Bei diesem Dedifferenzierungsvorgang scheint neben der Umschaltung der Synthese von Typ II auf Typ I auch die Expression der anderen Kollagentypen (1 α , 2 α , 3 α , kurzkettige Kollagene) gestoppt zu werden. Statt dessen findet sich die Synthese von Typ V Kollagen (43).

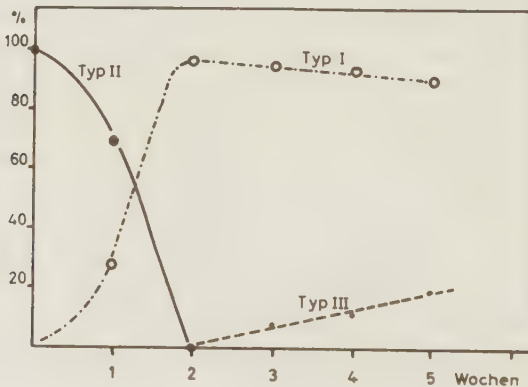


Abb. 4 : Veränderungen der Kollagensynthese bei der Dedifferenzierung von Chondrocyten aus Kalb (42)

Regulation auf der Ebene der Genexpression

Einige Veränderungen des Synthesemusters lassen sich auch auf der Ebene der Expression spezifischer Gene, der Synthese spezifischer mRNAs verfolgen. Durch die In Vitro Translation isolierter mRNA konnte gezeigt werden, daß der Gehalt an translatierbarer, für Kollagen Typ II spezifischer mRNA, während der Umschaltung der Chondrocyten abnahm, Typ I

kodierende mRNA andererseits neu gebildet wurde (44). Dieser Umschaltprozeß betrifft nicht nur die Kollagengene. Auch die mRNA der Linkproteine und des Coreproteins der Proteoglykane zeigen eine koordinierte Abnahme, während gleichzeitig die mRNAs von Typ I Kollagen und Fibronektin neu gebildet werden (45). Da diese alternativen Synthesemuster bei der In Vitro Translation teilweise nebeneinander vorliegen, in vivo aber nur eine ausschließliche Expression eines der beiden Muster beobachtet werden kann, spielen wahrscheinlich noch Regulationsvorgänge auf der translationalen Ebene eine Rolle (46). Dieser Umschaltvorgang auf der mRNA Ebene konnte auch für die beschleunigte Dedifferenzierung durch BrdU gezeigt werden (47). Alle diese Versuche beruhten auf dem Nachweis und der Mengenbestimmung translatierbarer mRNAs. Ein genauer Einblick in die Veränderung von mRNA Konzentrationen ist erst möglich durch die Verwendung von cDNA Klonen, die komplementären Sequenzen spezifischer mRNAs entsprechen, die in einem Plasmidvektor kloniert wurden. Diese Klone bilden den Ausgangspunkt sowohl zur Charakterisierung der mRNA-Struktur als auch zur Isolierung der kodierenden Gene aus einer Genbank.

Charakterisierung der Kollagengene

Für die beiden Peptidketten des Typ I Kollagens gelang es als erstes spezifische cDNA Klone zu konstruieren (48,49). Später wurden auch cDNA Klone für Kollagen Typ III (50) und Typ IV (51) isoliert. Mit Hilfe dieser Klone war es möglich, die Gene der Peptidketten von Typ I und Typ III Kollagen zu isolieren und zu charakterisieren (51). Alle diese Gene weisen eine besonders komplexe Exon-Intron Struktur auf. Die kodierenden Sequenzen sind auf etwa 50 Exons über einen Bereich von etwa 40 kb verteilt. Beim Spleißen des primären Transkripts werden alle Introns ausgeschnitten und es entsteht eine mRNA, die nur noch etwa 10% der Länge des primären Transkripts entspricht. Die Exons weisen im Bereich der Tripelhelix eine Länge von Vielfachen einer 54 bp Ein-

heit auf. Diese Repetition im Exonmuster deutet auf eine Entstehung des Gens durch eine Folge von Amplifikationen kurzer Ausgangssequenzen hin (53).

Lange Zeit standen für Matrixkomponenten des Knorpels keine spezifischen cDNA Klone zur Verfügung. Erst in neuester Zeit gelang es, cDNA Klone mit Kollagen Typ II spezifischen Sequenzen zu isolieren (54,55). Mit Hilfe dieser Klone konnte ein Teil des Typ II Gens von Huhn (56) und Mensch (57) isoliert werden. Der prinzipielle Aufbau unterscheidet sich, zumindest in den bisher charakterisierten Bereichen, nicht wesentlich von anderen Kollagengenen. Als Besonderheit scheint im Falle des Gens bei Huhn durch die Verkürzung einiger Introns, bei einer ähnlichen Anzahl von Exons, eine kompaktere Struktur erreicht worden zu sein.

Daneben konnten noch cDNA Klone isoliert werden, die zwei minoren Kollagentypen des Knorpels entsprechen (58). Für andere, nicht-kollagene Komponenten der extrazellulären Matrix wurden bisher noch keine spezifischen Klone isoliert. Solche Klone bieten erst die Möglichkeit zu einem weitergehenden Verständnis der grundlegenden, molekularen Regulationsvorgänge während der Chondrogenese und der Dedifferenzierung von Chondrocyten.

Charakterisierung knorpelspezifischer cDNA Klone

In dieser Arbeit wird die Untersuchung mehrerer cDNA Klone beschrieben, die mit der Zielsetzung konstruiert wurden, weitere knorpelspezifische Proben zu erhalten. Als Ausgangsmaterial zur Herstellung der cDNA Klone wurde angereicherte mRNA, isoliert aus *Sterna* embryonaler Hühner, mit einer Länge von mehr als 4 kb verwendet. In diesem Längenbereich finden sich neben der mRNA für Kollagen Typ II weitere mRNA Species hochmolekularer Matrixkomponenten (60). Die komplementäre DNA wurde ausgehend von einem oligo-dT Primer synthetisiert und danach in die doppelsträngige Form übergeführt. Die cDNA wurde nun nicht sofort in einen Vektor eingebaut, sondern mit verschiedenen Restriktionsenzymen ge-

spalten. Dabei zeigte das Auftreten majorer Fragmente die Existenz mengenmäßig vorherrschender mRNA Species an. Die Fragmente einer solchen Spaltung (mit Bam HI/ Hind III) wurde nun in den Vektor pBR 322 kloniert. Durch die Hybridisierung mit markierter mRNA des Knorpels konnte eine Reihe spezifischer Klone isoliert werden (59,44). Drei dieser Klone, p550, p1300 und p2200, werden in dieser Arbeit näher untersucht und charakterisiert.

Für einen dieser Klone, p550, konnte bereits gezeigt werden, daß komplementäre Sequenzen in zwei hochmolekularen mRNA Species (4 und 6 kb) zu finden sind, deren Synthese während der Dedifferenzierung von Chondrocyten abgeschaltet wird (44,59). Die Abnahme im Gehalt dieser mRNAs zeigt dabei sowohl einen parallelen Verlauf zur Abnahme der Menge an translatierbarer mRNA von Kollagen Typ II (44), als auch der Menge an synthetisiertem Typ II Kollagen (Abb. 4). Die Expression des p550 Gens schien demnach spezifisch für den differenzierten Knorpelzustand zu sein.

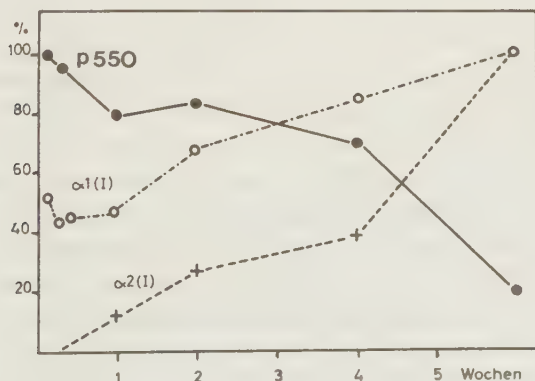


Abb. 5 : Veränderungen des mRNA Gehalts während der Dedifferenzierung von Chondrocyten aus Hühnersterna

führten deshalb zu der anfänglichen Vermutung, daß es sich um einen spezifischen Klon für Kollagen Typ II handeln mußte. Der Vergleich der DNA Sequenz mit bekannten Sequenzen des Kollagen Typ II und anderen Kollagenen bestätigte diese Hypothese jedoch nicht. Da keine weiteren Proteinsequenzen von Matrixkomponenten des Knorpels bekannt sind, blieb die Identität des Klons p550 deshalb bisher ungeklärt.

Ein weiterer Klon, p2200, zeigte im Gegensatz zu p550 eine Hybridisierung mit sehr viel größeren mRNA Species des Knorpels mit einer Länge von etwa 7 und 11 kb. Der Gehalt dieser mRNAs nimmt im Verlauf der Chondrogenese von Limb Buds stark zu und scheint ebenfalls eine typische Komponente des Knorpels zu repräsentieren. Größe und Art der Regulation führten zu der Vermutung, daß diese Sequenz für das Coreprotein eines Proteoglykans kodieren könnte (61).

Über den Klon p1300, der ebenfalls durch die Hybridisierung mit mRNA aus *Sterna* als knorpelspezifischer cDNA-Klon identifiziert worden war, lagen bisher keine weiteren Informationen vor.

Die Aufrechterhaltung der Struktur und damit der physiologischen Funktion des Knorpels hängt entscheidend von der koordinierten Expression der einzelnen Matrixkomponenten ab. Störungen in der Synthese und der Modifikation von Kollagenen und Proteoglykanen führen zu schwerwiegenden pathologischen Zuständen (62). Welche molekularen Ursachen dafür verantwortlich sind, läßt sich erst entscheiden, wenn ein Einblick in die grundlegenden Regulationsmechanismen solcher Gene möglich geworden ist. Einen Ansatzpunkt zur Beantwortung solcher Fragestellungen bietet die Untersuchung der vorhandenen knorpelspezifischen cDNA Klone, da mit deren Hilfe Aussagen über Proteinstruktur, Aufbau und Regulation bisher noch nicht beschriebener Komponenten gemacht werden können.

B. AUFGABENSTELLUNG

Die koordinierte Expression strukturell und funktionell unterschiedlicher Matrixkomponenten ist eine Voraussetzung zur Ausbildung funktionsfähiger Bindegewebe. Im Falle des Knorpels wurden bisher nur die mengenmäßig vorherrschenden Komponenten (Kollagen Typ II, Proteoglykane) in ihrer Proteinstruktur genauer untersucht. Über den Aufbau der minoren Komponenten der Knorpelmatrix (Glykoproteine, verschiedene Gruppen von Proteoglykanen) ist bisher noch sehr wenig bekannt, ebenso über Struktur und Regulation der kodierenden Gene (Ausnahme : Kollagen Typ II).

Einen Zugang zu einzelnen Komponenten bietet die Verwendung von cDNA-Klonen, die knorpelspezifisch exprimierten mRNAs entsprechen (p550, p2200 und p1300; s. Einleitung). Mit Hilfe dieser Klone sollte es möglich sein, sowohl die Expression in verschiedenen Geweben zu verfolgen, als auch die homologen Gene aus Genbanken zu isolieren und genauer zu untersuchen. Ziel der Untersuchungen war es damit, ein Bild von der Expression, Gen- und Proteinstruktur verschiedener Matrixkomponenten zu erhalten.

Im einzelnen ergaben sich damit folgende Aufgaben :

- (1) Charakterisierung der cDNA-Klone (p550, p1300 und p2200) durch Restriktionskartierung und Sequenzierung
- (2) Identifizierung der kodierten Genprodukte durch Sequenzvergleiche und 'Hybrid-Selektion-Translation'
- (3) Vergleich der Expression der homologen mRNAs in unterschiedlichen Geweben des Huhns
- (4) Konstruktion von repräsentativen Genbanken
- (5) Isolierung der entsprechenden Gene aus den Genbanken und Analyse ihrer Struktur

C. ERGEBNISSE

1. Charakterisierung des cDNA-Klons p550

Der cDNA-Klon p550 war in vorhergehenden Arbeiten (44,59) bereits teilweise charakterisiert worden und repräsentiert eine anscheinend typische Komponente des Knorpels. Eine Sequenzierung des Klons stand noch aus, ebenso wie eingehendere Versuche zur Identifizierung des Genprodukts und dessen Verteilung in verschiedenen Geweben. Im ersten Teil der Arbeit wird deshalb die Untersuchung dieses Klons genauer beschrieben. Da ein Großteil der verwendeten Methoden im Verlauf der weiteren Arbeit häufig angewandt wurde, werden sie exemplarisch in diesem Teil genauer beschrieben.

1.1 Klonierung von Fragmenten des cDNA-Klons p550

Die Restriktionskarte und die bereits sequenzierten Bereiche des cDNA Klons p550 sind in der Abb. 6 angegeben. Die terminalen Restriktionsstellen Bam HI und Hind III repräsentieren die zur Klonierung des doppelsträngigen cDNA Fragments in den Vektor pBR 322 verwendeten Spaltstellen. Mit der Methode von Maxam und Gilbert (64) wurde die Sequenz eines großen Teils des Inserts bestimmt. Nur ein Bereich von etwa 80 bp Länge in der Umgebung der Sal I Spaltstelle fehlte noch zu einer vollständigen Sequenz. Aus der Restriktionskarte läßt sich entnehmen, daß die noch fehlende Sequenz vollständig in einem Hpa II Fragment mit einer Länge von etwa 112 bp enthalten ist. Dieses Fragment wurde deshalb für die Klonierung ausgewählt, da die Sequenzen an den beiden Enden bereits bekannt waren und damit eine leichte Einordnung in die Gesamtsequenz ermöglichten. Zur Isolierung

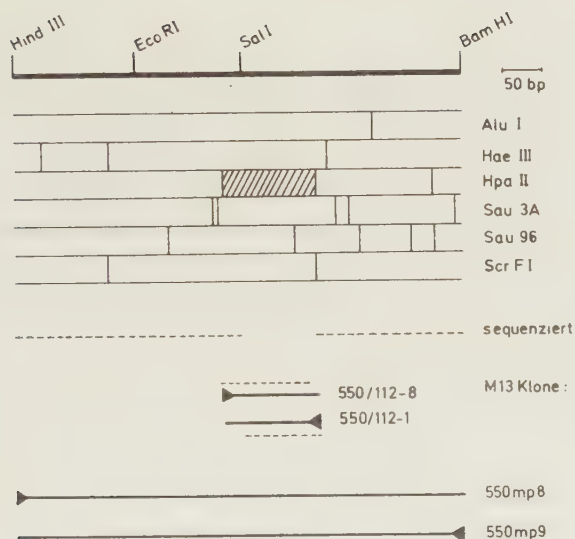


Abb. 6 : Restriktionskarte des cDNA Klon p550 (Insert)

/// : Subkloniertes Hpa II Fragment
 ► : Sequenzierungsrichtung der M13 Klone
 --- : Sequenzierte Bereiche

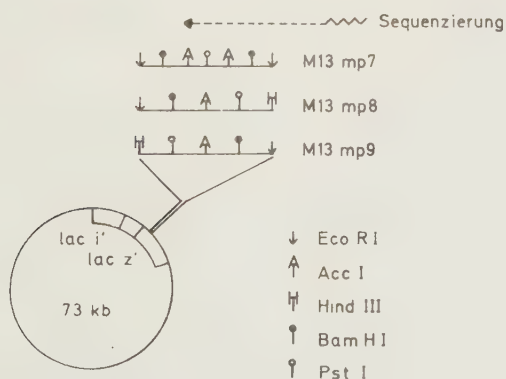


Abb. 7 : Restriktionskarten der verwendeten M13 Vektoren

~~~~~ : 15 b Sequenzierungsprimer  
 --> : Sequenzierungsrichtung

dieses Fragments wurde das Insert von p550 mit Hpa II (C/CGG) gespalten, um eine sichere Abtrennung des 112 bp Fragments von den Fragmenten des Vektors pBR 322 bei der Elektrophorese in einem Polyacrylamidgel zu erreichen.

Die Sequenz dieses Fragments wurde nicht direkt nach der Methode von Maxam & Gilbert bestimmt, sondern erst nach der Klonierung in dem Einzelstrangvektor M13 mp7 (65). Der Hauptvorteil dieses Vorgehens liegt in der Möglichkeit der technisch einfachen Sequenzierung des Fragments mit der Dideoxymethode nach Sanger (67). Dieses System bietet daneben noch eine Reihe weiterer Vorteile, da die DNA sowohl in der einzelsträngigen Phagenform als auch in der doppelsträngigen, replikativen Form isoliert, und so für unterschiedlichste Zwecke verwendet werden kann.

Als Vektor wurde M13 mp7 verwendet (66), der eine Anzahl von Restriktionsstellen ('Polylinker') innerhalb eines engen Bereichs des lac z' Gens enthält, die zum Einbau verschiedener Fragmente verwendet werden können (Abb. 7). Rekombinante Klone lassen sich deshalb nach der Transformation in einen geeigneten Empfängerstamm (JM 103) am Ausfall einer aktiven  $\beta$ -Galaktosidase erkennen. Durch die Integration eines Inserts innerhalb des lac z' Gens bildet der Phage ein Fusionsprotein, das nicht mehr in der Lage ist, durch  $\alpha$ -Komplementation mit dem lac zAM Peptid von JM 103 eine funktionelle  $\beta$ -Galaktosidase auszubilden. Die Aktivität des Enzyms läßt sich durch die Spaltung des farblosen X-Gal zu einem blauen Farbstoff nachweisen. Phagen mit eingebautem Insert erscheinen deshalb als weiße Plaques, im Gegensatz zu den blau gefärbten Plaques des Wildtyps. Die Plaques entstehen dabei nicht durch die Lyse der Bakterien, sondern durch verlangsamtes Wachstum der infizierten Zellen.

Die doppelsträngige DNA (replikative Form) des Vektors enthält zwei durch Acc I spaltbare Sequenzen (GT/CGAC) innerhalb der Polylinkerregion im lac z'-Gen. Diese Stellen können zur Klonierung des 112 bp Hpa II Fragments (C/CGG) verwendet werden, da in beiden Fällen die gleichen Einzelstrangenden (CG-) entstehen und damit miteinander ligierbar

sind. Bei der Spaltung des Vektors entsteht neben dem linearisierten Vektor ein zusätzliches sehr kurzes Fragment, das vor der Ligation nicht abgetrennt wurde. Das isolierte 112 bp Fragment (etwa 0.01 ng/ $\mu$ l) wurde mit dem gespaltenen Vektor (2 ng/ $\mu$ l) in einem Volumen von 10  $\mu$ l ligiert (s. G.2.9.). Das molare Verhältnis von Vektor (7.3 kb) zu Insert (0.1 kb) betrug dabei etwa 3:1. Zur Kontrolle der Transformation wurde Ligationspuffer ohne DNA und außerdem gespaltenen und ungespaltenen Vektor (je 10 ng/ $\mu$ l) transformiert.

Das Ergebnis der Transformation ist in der Tabelle 1 zusammengefaßt. Zur Transformation von JM 103 wurden aliquote Anteile der Ligationsansätze eingesetzt. Die Angaben in Tabelle 1 beziehen sich auf die gesamten Ligationsansätze.

Tabelle 1 : Klonierung des 112 bp Hpa II Fragmentes

| Ansätze                       | ng Vektor | Plaques |         | Plaques/ng |
|-------------------------------|-----------|---------|---------|------------|
|                               |           | weiß    | blau    |            |
| Kontrolle (keine DNA)         | 0         | 0       | 0       | 0          |
| M13 mp7/ ungespalten          | 100       | 0       | (20000) | (200)      |
| M13 mp7/ Acc I                | 100       | 0       | 60      | <1         |
| M13 mp7/ Acc I/<br>Religation | 100       | 0       | (6000)  | (60)       |
| M13 mp7 + 112 bp Hpa II       | 20        | 22      | 540     | 28         |

Bei der Ligation des 112 bp Fragmentes fanden sich 22 positive weiße Plaques neben 540 blauen Plaques; das entspricht etwa 4% aller Klone. Diese kleine Ausbeute an rekombinanten Klonen erklärt sich aus der geringen Menge an eingesetztem Insert gegenüber den vorhandenen Vektormole-

külen (Vektor : Insert = 3:1). Daneben kann das kleine, bei der Acc I Spaltung des Vektors entstehende Fragment ebenfalls eingebaut werden und ist bei der Ligation begünstigt, da es in größerer Anzahl vorliegt (Acc-Fragment : 112 bp Hpa II = 3:1). Der Einbau dieses Fragments stört den Leseraster des lac z' Peptids nicht, weshalb in diesem Fall blaue Plaques entstehen. Weiße Plaques können damit nur rekombinanten Klonen entsprechen. Die mögliche Zirkularisierung des Vektors ohne Einbau eines Fragments ergibt ebenfalls nur blaue Plaques.

Die Effizienz der Transformation läßt sich aus der Ausbeute bei der Transformation des ungespaltenen Vektors erkennen. Etwa 200 Plaques/ng Vektor wurden bei diesem Versuch erreicht. Bei der Transformation des gespaltenen und religierten Vektors finden sich etwa 30% (60/200) der bei Ligation aller Moleküle erreichbaren Anzahl wieder. Die Spaltung von M13 mp7 mit Acc I erfolgte vollständig, da weniger als 1 Plaque/ng durch die Transformation von gespaltenem Vektor auftrat. Erwartungsgemäß fanden sich auch bei der Transformation des Ligationspuffers keine Plaques. Die Ligations- und Transformationsausbeuten zeigten deutliche Schwankungen, da in anderen Versuchen Werte bis zu 1000 Plaques/ng Vektor erreicht werden konnten.

Von den erhaltenen 22 weißen Klonen, die das 112 bp Hpa II Fragment eingebaut enthalten sollten, wurden 10 willkürlich ausgewählt (550/112-1 bis -10) und genauer charakterisiert. Trotz der geringen Längendifferenz zwischen Vektor (7.3 kb) und rekombinantem Klon (7.4 kb) zeigte die DNA der ausgewählten Klone eine reduzierte Wanderungsgeschwindigkeit bei der Gelelektrophorese gegenüber dem Vektor. Einen besseren Hinweis auf den Einbau eines zusätzlichen Fragments gab jedoch der folgende Versuch.

Da das klonierte Fragment an beiden Enden eine Hpa II Stelle aufwies, konnte es im Vektor in zwei verschiedenen Orientierungen eingebaut werden. Der Replikationsmechanismus des M13 Phagen führt jedoch zur Auswahl eines bestimmten Strangs - des (+)-Strangs - bei der Synthese der einzel-

strangigen Phagen-DNA (65). Enthalten zwei rekombinante Klone ein Fragment in unterschiedlicher Orientierung, finden sich deshalb in der Phagen-DNA die Inserts als komplementäre Sequenzen. Dieser Effekt läßt sich zur Bestimmung der relativen Orientierung ausnutzen. Hybridisiert man die einzelstrangigen Phagen-DNA's zweier Klone miteinander, können nur komplementäre Insertbereiche miteinander paaren; es entsteht ein partiell doppelsträngiges Molekül, das bei der Gelelektrophorese als zusätzliche, langsamer wandernde Bande erkannt werden kann. Das Fehlen einer solchen Bande zeigt die Parallelität des Inserts zweier Klone an. Auf diese Weise ließ sich zeigen, daß die meisten Klone, mit Ausnahme von 550/112-7, -8 und -9, die gleiche Orientierung aufwiesen. Zur Sequenzierung wurden deshalb die beiden Klone 550/112-1 und -8 ausgewählt. Die komplementäre Anordnung der Inserts wurde dabei bestätigt (s.C.1.2.).

Neben dem 112 bp Hpa II Fragment wurde prinzipiell auf die gleiche Weise das gesamte Insert von p550 in M13 rekloniert. Um die Einbaurichtung dieses Bam HI/ Hind III-Fragments definieren zu können, wurden die beiden Vektoren M13 mp8 und mp9 verwendet (Abb. 7). Die Vektoren wurden ebenfalls mit Bam HI und Hind III gespalten, mit dem isolierten Insert von p550 ligiert und in JM 103 transformiert. Alle getesteten weißen Plaques zeigten bei der Gelelektrophorese den Einbau eines zusätzlichen Fragments an. Einzelne Klone wurden willkürlich ausgewählt und werden im Folgenden als 550mp8 und 550mp9 bezeichnet.

In Abb. 6 sind Lage und Orientierung der M13 Subklone des cDNA Klones p550 zusammenfassend angegeben.

## 1.2 Sequenzierung der Klone 550/112-1 und -8

Zur Sequenzierung der eingebauten Fragmente von 550/112-1 und -8 wurde die 'Didesoxymethode' nach Sanger verwendet (67). Das Prinzip ist in Abb. 8 am Beispiel der Sequenz des Klon 550/112-8 schematisch zusammengefaßt. In einem ersten Schritt wird ein kurzer einzelsträngiger, synthetischer Oligonukleotidprimer (15 b: s.G.2.10.) an die einzelsträngige Phagen-DNA gebunden. Dieser Primer bindet etwa 10 b vor dem Bereich der Klonierungsstellen des Vektors und die DNA-Polymerase (Klenow Fragment) kann davon ausgehend einen komplementären Strang synthetisieren, der in den Bereich des klonierten Inserts hinein verlängert wird.

Die für eine Sequenzierung nötigen basenspezifischen Kettenabbrüche lassen sich durch die Verwendung von 2',3'-Didesoxynukleotiden (ddNTP) erreichen, da nach deren sequenzabhängigen Einbau eine weitere Verlängerung nicht mehr möglich ist. Es entsteht damit eine Anzahl verschieden langer Fragmente mit identischen 5'-Enden (=Primer) und einem eingebauten, sequenzspezifischen ddNTP am 3'-Ende. Durch den Einbau eines radioaktiv markierten Desoxynukleotids (meist  $\alpha^{32}\text{P}$ -dATP) lassen sich diese neu synthetisierten Stränge spezifisch markieren. Nach der Denaturierung der partiell doppelsträngigen Moleküle und deren längenabhängiger Auftrennung in Sequenzgelen, lassen sie sich deshalb durch Autoradiographie sichtbar machen.

Das Verhältnis der Didesoxynukleotide und der entsprechenden Desoxynukleotide in einem Ansatz ist entscheidend für die Länge der entstehenden Fragmente. Folgende molare Verhältnisse wurden verwendet (s.G.2.10.) :

ddA : dA = 45

ddG : dG = 40

ddT : dT = 70

ddC : dC = 40

Konzentration der dNTP = 1.66  $\mu\text{M}$

Jedes entstehende Fragment enthält neben dem Primer (15

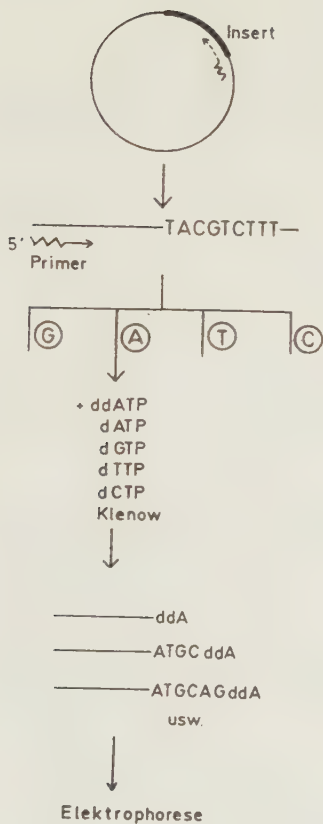


Abb. 8 :  
Prinzip der Sequen-  
zierung nach der Di-  
deoxymethode (67)  
Beispiel : 550/112-8  
Pos. 266-304

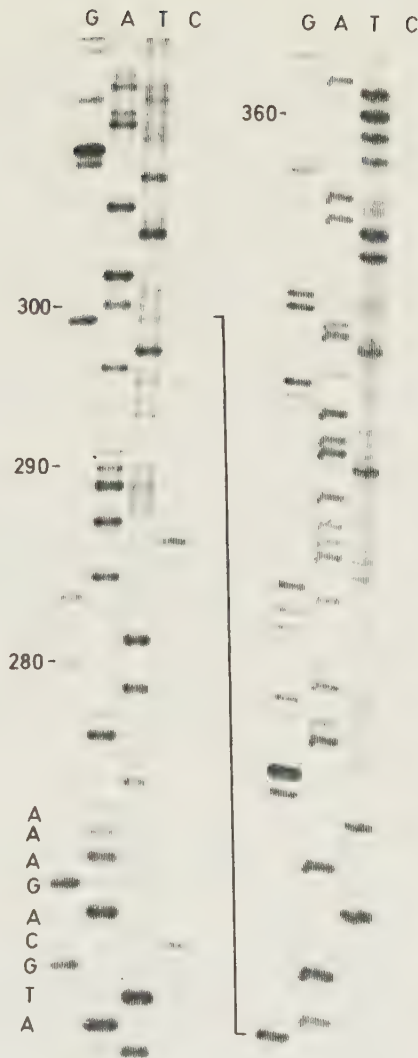


Abb. 9 :  
Sequenzierung des  
Klons 550/112-8  
Position : 266-360 (Abb. 10)  
Sequenz s. Abb.10  
Gel : 8% Acrylamid  
50-55°C

b) auch noch den Bereich bis zum Beginn der klonierten Sequenz (10-20 b, je nach Klonierungsstelle) und ist somit mindestens etwa 25 b lang. Zur Auftrennung dieser Fragmente reichen deshalb 6-8% Acrylamidgele aus.

Die Autoradiographie eines solchen Sequenzgels ist in Abb. 9 für die Anfangssequenz des Klons 550/112-8 gezeigt. Die lesbare Sequenz beginnt hier bei der Position 8 des klonierten Fragments und kann bis etwa Position 115 sicher gelesen werden. Die Sequenz des Klons 550/112-1 wurde auf die gleiche Weise bestimmt, wobei etwa 100 Basen sicher gelesen werden konnten.

Abb. 10 : Überlappende Sequenzen von 550/112-1, 550/112-8  
mit den Randsequenzen von p550 (59)  
::: : Komplementarität der Sequenzen  
CCGG : Hpa II Klonierungsstellen  
Nummerierung entspricht Abb. 11

```

                251
p550      5'- TCAACCGGACGGATTATGCAGAAATC
112/8      5'- TTATGCAGAAATCAATGTCGACCACAAATTCC
                ::::::::::::::::::::
112/1      3'- CAGCTGGTGTTTAAGG

                301
p550      ATCGACACCTCATTGGCAAGAACGGAGCAAACATTAACAGGATTA
112/8      ::::::::::::::::::::
112/1      TAGCTGTGGAGTAACCGTTCTTGCTCGTTTGTAAATTGCTCTAAT

                351                                380
p550      AGGACCTCTACAAGGTCTCTGTACGCATT - 3'
112/8      ::::::::::::::::::::
112/1      TCCTGGAGATGTTCCAGAGACATGCGTAAGG - 5'

```

Da die beiden Klone das 112 bp Fragment in entgegengesetzter Orientierung enthalten, müssen die beiden Sequenzen überlappen. In der Abb. 10 sind die gefundenen Sequenzen von 550/112-1 und -8 miteinander und mit den bereits bekannten Randsequenzen (59) zur Deckung gebracht worden. Da sich weder Unstimmigkeiten in den komplementären Sequenzen noch im

Vergleich mit den Randsequenzen ergaben, ließ sich eine eindeutige Gesamtsequenz ableiten, die in Abb. 11 gezeigt ist. Die Länge des p550 Inserts beträgt damit 547 bp, die Länge des klonierten Hpa II Fragments 118 bp.

### 1.3 Orientierung und Proteinsequenz des cDNA Klonen p550

#### Gesamtsequenz und hypothetischer Leseraster von p550

Die Gesamtsequenz des klonierten cDNA Fragmentes in p550 ist in Abb. 11 angegeben und repräsentiert damit die (komplementäre) Sequenz einer bestimmten mRNA-Species. Falls das klonierte Fragment ausschließlich proteinkodierende Bereiche enthält, müßte sich ein durchgehend lesbarer Raster ohne Stopcodons finden. Sind dagegen 5'- oder 3'-terminale, nichtkodierende Bereiche enthalten, ist die Identifizierung des kodierenden Strangs nicht mehr so einfach möglich.

Die Sequenz von p550 wurde deshalb auf durchgehend lesbare Leseraster hin abgesucht. Unter den sechs möglichen Rastern fand sich nur einer, der sich in ein durchgehendes Protein übersetzen ließ. Alle anderen Raster enthielten meist mehrere Stopcodons. Die Sequenz des einzigen vollständig translatierbaren Rasters beginnt mit der Sequenz AAG-CTT-GGG-. Die davon abgeleitete Proteinsequenz mit 182 Aminosäuren und die Aminosäurezusammensetzung dieses Teilfragments des p550 Genprodukts ist ebenfalls in der Abb. 11 angegeben. Für die Orientierung der Sequenz auf der mRNA ergibt sich damit folgende Anordnung: 5'-Hind III-Bam HI-3'. Ob dieser Leseraster tatsächlich verwendet wird, läßt sich aus diesen Daten alleine noch nicht ableiten. Erst die Aussage über die tatsächliche Orientierung dieser Sequenz auf der mRNA macht den gefundenen hypothetischen Leseraster wahrscheinlicher.

Abb. 12 : DNA- und abgeleitete Proteinsequenz von p550  
Aminosäure unter der ersten Base eines Codons

```

      10              30              50
AAGCTTGGGCAAGCATTGA•CTGAAGTCTAT•GCAAAGGCCAACAGTTT•TACCGTCTCCTCG•
K L G Q A L T E V Y A K A N S F T V S S

      70              90              110
GTCTCTGCC•CCCTCTTGGC•TTTCATCGTTT•CATTATTGGTTT•TGTTGGTCA•AAACCTGGCC•
V S A P S W L H R F I I G L F G Q N L A

      130              150              170
AAAATAACTCAGCAAATGCC•AAAGATTCA•CATCGAATTCA•CTGAAGGAGAGGATAAAATC•
K I T Q Q M P K I H I E F T E G E D K I

      190              210              230
ACTTCGGAAGGACCTACAGA•AGATGTGAATGTGGCTCAGGAACAGATTGAAGTCATGGTC•
T S E G P T E D V N V A Q E Q I E V M V

      250              270              290
AAGGATCTGATCAACCGGAC•GGATTATGCAGAAATCAATGTCGACCACAAATTCCATCGA•
K D L I N R T D Y A E I N V D H K F H R

      310              330              350
CACCTCATTGGCAAGA•ACGGAGCAAACATTAA•CAGGATTAAAGGACCTCT•ACAAGGTCTCT•
H L I G K N G A N I N R I K D L Y K V S

      370              390              410
GTACGCATTCCCCCGGACA•ATGAGAAGAGCAACCTGATCAGAATTGAAGGAGATCCACAG•
V R I P P D N E K S N L I R I E G D P Q

      430              450              470
GGGGTCCAAACAGGCCAAGA•AAAGAGCTGCTGGA•ACTCGCTTCCCGTATGG•AAAATGAACGC•
G V Q Q A K K E L L E L A S R M E N E R

      490              510              530
ACCAAGGACCTAATCATTGAGCAGAAAT•TCCACCGGACCATCATTGGGCAGAAGGGTGAG•
T K D L I I E Q K F H R T I I G Q K G E

      547
CGGATCC
R I

```

Abgeleitete Aminosäurezusammensetzung des obigen  
Teilfragments des p550 Proteins (182 Aminosäuren)

|        |        |         |        |
|--------|--------|---------|--------|
| Asp 50 | Glu 60 | Val 60  | Phe 33 |
| Asn 60 | Pro 33 | Met 16  | His 33 |
| Thr 50 | Gly 60 | Ile 110 | Lys 88 |
| Ser 50 | Ala 55 | Leu 71  | Arg 55 |
| Glu 93 | Cys 0  | Tyr 16  | Trp 5  |

(Angaben in Aminosäuren/ 1000)

## Bestimmung der Orientierung der p550 Sequenz

Zur Bestimmung der Orientierung läßt sich die Tatsache ausnutzen, daß nur einer der beiden Stränge der cDNA Sequenz komplementär zur mRNA sein kann und deshalb mit mRNA hybridisieren kann. Die einfachste Möglichkeit, strangspezifische Hybridisierungsproben zu erhalten, bietet wiederum die Verwendung von M13 Klonen.

Das gesamte Insert des Klons p550 war bereits in zwei unterschiedlichen Orientierungen in die Vektoren M13 mp8 und mp9 eingebaut worden (s.C.1.1.). Die Klone 550 mp8 und 550 mp9 enthalten in der einzelsträngigen Phagenform nur jeweils einen der beiden komplementären Stränge des Inserts. Aus diesem Grund kann nur einer der beiden Klone mit der entsprechenden mRNA hybridisieren. Die Orientierung läßt sich somit aus der Einbaurichtung des Fragments ableiten.

Die für die Hybridisierung notwendigen radioaktiv markierten Proben lassen sich auf einfache Weise nach der Methode von Hu & Messing (69) herstellen. Ein kurzer einzelsträngiger Primer (17 b; s.G.2.5.) wird an die einzelsträngige Phagen-DNA gebunden, der - anders als bei der Sequenzierung - 5'-terminal 'nach' dem klonierten Insert bindet. Verlängert man nun diesen Primer mit DNA-Polymerase/Klenow Fragment unter Zugabe aller Deoxynukleotide, wobei eines radioaktiv markiert vorliegt (meist  $\alpha^{32}\text{P}$ -dCTP), erhält man ein nur im Bereich des Vektors partiell doppelsträngiges Molekül. Die Sequenz des klonierten Inserts befindet sich dagegen in den einzelsträngigen Anteilen des Moleküls (Abb. 12B). Der partiell doppelsträngige, nur in einem Strang markierte Bereich ist meist etwa 2-4 kb lang und umfaßt damit ausschließlich Teile des Vektors (Gesamtlänge etwa 7.3 kb). Diese markierten Moleküle können nun direkt zur Hybridisierung verwendet werden.

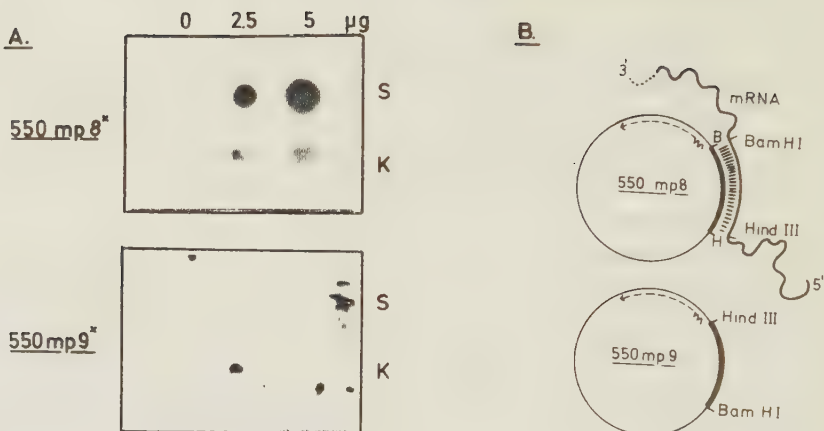
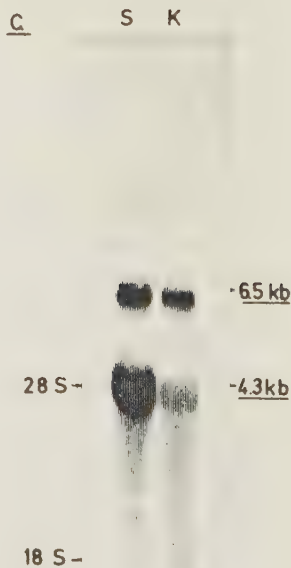


Abb. 12 :  
Bestimmung der Leserichtung  
der p550 Sequenz

- A. Dothybridisierung von markiertem 550mp8\* und 550mp9\* mit verschiedenen Mengen RNA
- B. Theoretische Ableitung der Orientierung der p550 Sequenz auf der mRNA  
~~~~~ : Primer und neu synthetisierter, markierter Komplementärstrang
- C. Northernhybridisierung mit markiertem 550mp8*
10 μ g RNA wurde mit Glyoxal denaturiert, auf einem 1% Agarosegel aufgetrennt, transferiert und mit markiertem 550 mp8* hybridisiert

RNA : S = Sterna / 17 Tage
K = Kalvarien/ "



Dot-Hybridisierung

Beide Klone - 550 mp8 und 550 mp9 - wurden nach dieser Methode markiert, wobei spezifische Aktivitäten von etwa 0.5×10^8 cpm/ μ g DNA erreicht werden konnten. Durch Dot-Hybridisierung wurde in einem ersten Versuch getestet, welcher der beiden Klone mit der RNA hybridisiert. Da die RNA sehr viele Sekundärstrukturen ausbilden kann, muß sie vollständig denaturiert werden, bevor sie an Nitrocellulose oder eine modifizierte Transfermembran gebunden und für die Hybridisierung verwendet werden kann. Dies wurde erreicht durch die Behandlung mit 1M Glyoxal bei 50°C (70), wobei durch die kovalente Bindung des Glyoxals an die Guaninreste der RNA alle Sekundärstrukturen aufgelöst werden.

Zur Hybridisierung wurde die RNA aus Sterna und Kalvarien von 17 Tage alten Hühnerembryonen verwendet. Verschiedene Mengen (0, 2.5, 5 μ g) wurden denaturiert, in parallelen Ansätzen auf die Filter aufgetropft, getrocknet und durch Backen (2 Std/ 90°C) stabil fixiert.

Die beiden Filter wurden mit 1×10^6 cpm der radioaktiv markierten Klone 550 mp8 und 550 mp9 in einem kleinen Volumen (etwa 200 μ l/cm²) hybridisiert. Sowohl die Prähybridisierung als auch die eigentliche Hybridisierung erfolgte dabei bei 65 °C in 5x SSC (s.G.2.5.), dem 1% Sarkosyl NL 30 zugegeben worden war. Diese Methode (unter Verwendung des Detergens Sarkosyl NL 30 (71)), zeichnet sich gegenüber der üblichen Formamid-Methode (108) durch größere Einfachheit und eine verbesserte Sensitivität aus. Vorversuche zeigten, daß die Spezifität der Hybridisierung jedoch vergleichbar ist und die Methode sowohl bei DNA-DNA als auch bei DNA-RNA Hybridisierungen gute Ergebnisse liefert (75). Nach 20 Stunden wurde die nichtgebundene Probe durch Waschen mit 3x SSC/ 0.5% Sarkosyl und danach 3x SSC bei 65°C entfernt; die Filter wurden danach getrocknet und exponiert.

Das Ergebnis einer solchen Hybridisierung ist in Abb. 12A gezeigt. Eine Hybridisierung der RNA zeigte sich nur mit 550 mp8, nicht jedoch mit 550 mp9. Da die Orientierung des In-

serts in diesen beiden M13 Klonen festgelegt ist, läßt sich daraus die Orientierung der Sequenz erschließen, wie es in Abb. 12C gezeigt ist. Auf der mRNA findet sich damit folgende Anordnung : 5'-Hind III-Bam HI-3'.

Diese Orientierung entspricht der Richtung des gefundenen durchgehenden Leserasters (s. Abb. 11), dessen Verwendung in vivo damit sehr wahrscheinlich ist.

Identität der Proteinsequenz

Die abgeleitete Proteinsequenz des cDNA Klons p550 wurde nun verwendet, um sie mit bekannten Sequenzen in einer Datenbank (72) zu vergleichen. Es fand sich dabei keinerlei Ähnlichkeit mit einer bereits beschriebenen Proteinsequenz, sodaß auf diesem Wege eine Identifizierung des p550 Genprodukts nicht möglich war. Ausgeschlossen werden konnte zumindest eine Ähnlichkeit mit bekannten Kollagensequenzen. Auch der Vergleich der DNA-Sequenz mit beschriebenen DNA-Sequenzen führte zu keiner Identifizierung.

Längenbestimmung der p550 spezifischen mRNA

Die Dot-Hybridisierung ließ erkennen, daß nicht nur die RNA aus *Sterna*, sondern auch - in geringerem Maße - die RNA der Kalvarien mit 550 mp8 hybridisierte. Die Frage, ob die Hybridisierung mit einer bestimmten spezifischen mRNA Species erfolgte, oder ob eine unspezifische Bindung an die filtergebundene RNA stattfand, konnte durch Northernblotting und Hybridisierung mit 550 mp8 beantwortet werden. Die RNA von *Sterna* und Kalvarien wurde dazu wieder mit Glyoxal denaturiert und in einem 1% Agarosegel aufgetrennt (s.G.2.4.). Die RNA Moleküle trennten sich unter diesen Bedingungen ihrer Größe entsprechend auf. Nach der Auftrennung im Gel ließ sich die RNA durch Elektrottransfer annähernd quantitativ auf eine Transfermembran übertragen. Nach dem Trocknen und Backen der Filter wurden parallele Ansätze mit der markierten DNA von 550 mp8 und 550 mp9 hybridisiert.

Das Ergebnis der Dothybridisierung wurde bestätigt, da wiederum ausschließlich 550 mp8 zu einer Hybridisierung führte. In Abb. 12C kann man erkennen, daß zwei mRNA Species p550-spezifische Sequenzen enthalten. Die Längenbestimmung dieser RNAs erfolgte durch den Vergleich mit Glyoxal denaturierten DNA-Markern, die unter den verwendeten Bedingungen ein mit RNA-Fragmenten vergleichbares Laufverhalten zeigen (73). Die deutlich hybridisierende 4.3 kb Bande entspricht in der Länge der 28 S rRNA, was auch die Verbreiterung der Bande erklärt. Die zweite, schwächer hybridisierende mRNA hat eine Länge von etwa 6.5 kb.

Die Mengenverhältnisse der mRNA Species lassen sich durch Densitometrie der Autoradiographien grob abschätzen. Im Gegensatz zur RNA aus *Sterna* findet sich in der gleichen Menge RNA aus Kalvarien etwa 50% der 6.5 kb mRNA, aber nur etwa 5-10% der 4.3 kb mRNA. Die 4.3 kb mRNA repräsentiert in der *Sterna* RNA die vorherrschende Form, während in den Kalvarien die 6.5 kb mRNA in größerer Menge auftritt.

Die Existenz zweier unterschiedlich langer mRNAs wirft einige prinzipielle Fragen auf: welcher Zusammenhang besteht zwischen den beiden Formen? Welches ist die in vivo translatierte Form? Kodieren beide Formen das gleiche Protein? Diese Frage werden später im Zusammenhang mit der Genstruktur genauer diskutiert (s.D.5.).

1.4 Gewebespezifität des Klon p550

Die bisherigen Hybridisierungen haben gezeigt, daß das p550 Gen nicht ausschließlich im Knorpel des Sternums exprimiert wird, sondern auch in den Kalvarien, wenngleich in unterschiedlichen Mengen. Da sich die Matrixkomponenten von Kalvarien und Sternum deutlich unterscheiden, war bisher angenommen worden, daß keine Gemeinsamkeiten zwischen diesen beiden Geweben bestehen (74). Damit stellte sich die Frage, ob das p550-Gen eine gewebespezifische Expression zeigt, oder ob es generell exprimiert wird.

Die qualitative Beantwortung einer solchen Fragestellung ist möglich durch Northernblotting der RNA verschiedener Gewebe und deren Hybridisierung mit der p550-Sequenz. Dazu wurde nicht nur die RNA des hyalinen Knorpels der Sterna isoliert, sondern auch des epiphysealen Knorpels der Knochen (Fibula und Tibia). Bei 12 Tage alten Hühnerembryonen bestehen Fibula und Tibia noch größtenteils aus Knorpel, während bei 19 Tage alten Embryonen nur mehr die Endregionen aus dem epiphysenknorpel und dem artikulären Knorpel der Gelenke gebildet wird (im folgenden als 'Epiphysen' bezeichnet). Als knöchernes Gewebe, im Vergleich zu Kalvarien, wurde die RNA aus dem Mittelteil von Femur und Tibia (19 Tage alte Embryonen) isoliert, der keine knorpeligen Anteile mehr enthielt (im folgenden: 'Röhrenknochen'). Daneben wurde die RNA auch aus den davon völlig unterschiedlichen Geweben von Leber und Hirn extrahiert. Als weiteres Beispiel eines extrem differenzierten Bindegewebes dienten die Sehnen. In den meisten Fällen wurde die RNA aus den Geweben von 12, 17 und 19 Tage bebruteten Hühnerembryonen gewonnen. Ein sehr früher Entwicklungszustand wird durch kultivierte Embryofibroblasten repräsentiert, die von zwei Tage alten Embryonen stammen.

In allen Fällen wurde die RNA nach der Guanidinium-Methode (s.G.2.4.) aus den ganzen Geweben oder den kultivierten Zellen isoliert. Auf eine Anreicherung der mRNA durch Chromatographie an Oligo-dT Cellulose wurde verzichtet, um möglichst quantitative Aussagen über den mRNA Gehalt machen zu können. Gleiche Mengen dieser RNA-Präparationen (10 µg) wurden mit Glyoxal denaturiert, in einem 1% Agarosegel aufgetrennt, auf Filter transferiert und mit markiertem 550 mp8 hybridisiert. Das Ergebnis einer solchen Hybridisierung ist in Abb. 13 gezeigt. Hybridisierende mRNAs fanden sich in Sterna, Kalvarien und überraschenderweise auch in den Sehnen. Keine hybridisierenden mRNAs dagegen enthalten die RNAs von Leber, Hirn, Röhrenknochen und Embryofibroblasten.

In allen Fällen lassen sich wieder die bereits früher gefundenen mRNAs von 6.5 und 4.3 kb Länge nachweisen. Vergleicht man die Mengen in den verschiedenen Entwicklungs-

stufen, so scheinen im Sternum und in der Sehne keine Veränderungen im mRNA Gehalt zwischen Tag 12 und 19 der Embryonalentwicklung aufzutreten. Die Kalvarien zeigen jedoch eine

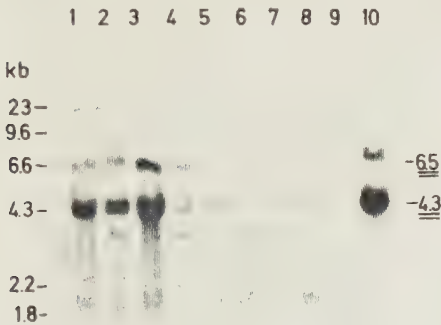


Abb. 13 :
Northernhybridisierung
mit markiertem 550 mp8*

Je 10 µg RNA wurden
mit Glyoxal denaturiert, in
einem 1% Agarosegel aufge-
trennt, transferiert und mit
markiertem 550 mp8 hybridisiert
(Autoradiographie)

- 1 = Sterna/ 12 Tage
- 2 = Sterna/ 19 "
- 3 = Kalvarien/ 12 "
- 4 = Kalvarien/ 19 "
- 5 = Embryofibroblasten
- 6 = Leber/ 16 "
- 7 = Gehirn/ 16 "
- 8 = Epiphysenknorpel/ 19 "
- 9 = Knochen/ 19 "
- 10 = Sehne/ 16 "

Tabelle 2 : Relative Expression des p550 Gens in
verschiedenen Geweben

| Stadium
mRNA (kb) | 12 Tage | | 19 Tage | |
|--------------------------------|---------|------|---------|-----|
| | 4.3 | 6.5 | 4.3 | 6.5 |
| Sternum | ++ | + | ++ | + |
| Kalvarien | +++ | ++ | + | + |
| Sehne | +++ | ++ | +++ | ++ |
| Röhrenknochen | | | | - |
| Epiphysen | | (+?) | | - |
| Leber (17 Tage) | | | | - |
| Hirn (17 Tage) | | | | - |
| Embryofibroblasten
(2 Tage) | | | | - |

deutliche Abnahme des mRNA Gehalts innerhalb dieser Zeitspanne, wobei am Tag 12 sogar mehr hybridisierende mRNA vorhanden ist als in den Sterna. Diese Abnahme geht außerdem einher mit einer Umkehr im Mengenverhältnis der beiden mRNA Species zueinander. Im Gegensatz zu allen anderen Geweben findet sich am Tag 19 hauptsächlich die 6.5 kb mRNA.

Im epiphysealen Knorpel der Röhrenknochen von 12 Tage alten Embryonen fanden sich nur sehr geringe Mengen der beiden mRNA-Species, die sich in den späteren Stufen jedoch nicht mehr nachweisen ließen.

In der Tabelle 2 sind die Mengen der mRNAs, die p550-spezifische Sequenzen enthalten, in einer grob-quantitativen Abschätzung angegeben. Diese Aussage über die relative Expression des p550 Gens in den verschiedenen Geweben konnte bei einer quantitativen Bestimmung bestätigt werden (75).

Dieses Ergebnis deutet darauf hin, daß der cDNA Klon p550 ein Gen repräsentiert, das nur in bestimmten Bindegeweben und, zumindest in Kalvarien, nur in bestimmten Entwicklungsphasen exprimiert wird. Diese Expression ist jedoch nicht gekoppelt an einen definierten Gewebetyp. Man findet die mRNA im Knorpel des Sternums, nicht aber im Gelenkknorpel, in den knöchernen Kalvarien, dagegen nur in geringen Mengen in langen Röhrenknochen. Außerdem sind die p550 spezifischen mRNA's in großer Menge in den Sehnen exprimiert, während in den Embryofibroblasten diese mRNAs nicht nachzuweisen sind. Die gefundene Verteilung der Expression, die nicht auf knorpelige Gewebe beschränkt ist, deutet auf eine funktionelle Bedeutung des p550 Genprodukts in flexiblen und elastischen Bindegeweben hin (s.D.1).

1.5 Direkte Identifizierung des p550 Proteins durch Hybrid Selektion

Die bisher beschriebenen Versuche erlaubten nur Aussagen über die Gewebespezifität der Expression und die Längen der kodierten mRNAs. Ein Rückschluß auf die Größe des transla-

tierbaren Proteins ist damit noch nicht möglich, da nicht notwendigerweise alle Bereiche der mRNA translatiert werden müssen.

Ließen sich die beiden, mit p550 hybridisierenden mRNA Species vollständig in Protein übersetzen, ergäbe das für die 6.5 kb mRNA ein Protein mit etwa 2200 Aminosäuren und einem Molekulargewicht von etwa 240 kd (bei etwa 110 d / Aminosäure). Die 4.3 kb mRNA ergäbe ein Protein von etwa 160 kd. Nimmt man an, daß etwa 80% der mRNA übersetzt werden, sollten sich die Proteine in einem Bereich von etwa 130 - 180 kd finden lassen.

Eine direkte Bestimmung der Größe des p550 Proteins ist möglich durch die Isolierung der spezifischen mRNA mit Hilfe der 'Hybridselektion' und deren anschließende Übersetzung in das zugehörige Protein durch 'In Vitro Translation'. Die selektive Anreicherung einer bestimmten mRNA-Species erreicht man durch Hybridisierung von RNA an einer trägergebundenen Hybridisierungsprobe. Die spezifisch gebundenen mRNA Moleküle lassen sich dadurch von allen anderen abtrennen und nach deren Freisetzen translatieren.

Zur Immobilisierung eignet sich besonders die kovalente Kopplung eines DNA-Fragments an DBM-Pulver (76). Die stabile Kopplung der Probe, das nötige geringe Volumen bei der Hybridisierung und die leichte Abtrennbarkeit des Pulvers von der Hybridisierungslösung machen dieses Verfahren vorteilhafter als die Kopplung einer Probe an Nitrocellulosefilter vor der Selektion (77).

Zur Anreicherung der p550 spezifischen mRNA wurde die einzelsträngige Phagen DNA von 550 mp8 kovalent an DBM-Pulver gebunden. Damit sollte eine höhere Ausbeute bei der Selektion erreicht werden, da im Gegensatz zu einer doppelsträngigen Probe die DNA vor der Bindung an den Träger nicht mehr denaturiert werden muß und vor allen Dingen ausschließlich komplementäre Sequenzen enthält.

Vor der Kopplung der DNA wurden die chemisch stabilen Aminogruppen von ABM-Cellulose Pulver (10 mg), durch die Behandlung mit NaNO_2 in saurer Lösung in Diazogruppen (DBM =

Diazobenzoxymethyl-) überführt. Etwa 20 μ g 550 mp8 DNA wurde sofort danach zugegeben. Nach der Bindung der DNA wurde in verschiedenen Waschschritten nicht oder nur lose gebundene DNA entfernt. In einem abschließenden Waschschriff (99% Formamid bei 65°C) sollten alle nicht kovalent gebundenen Moleküle abgelöst werden (s.G.2.11.). Dieses Pulver mit kovalent gebundener 550 mp8 DNA wurde in den folgenden Versuchen zur Hybridisierung verwendet.

Verschiedene RNA-Präparationen wurden mit der gebundenen Probe hybridisiert. Aus zwei Gründen wurden dazu keine angereicherten mRNA Fraktionen, sondern die gesamte RNA verwendet: erstens weist die Anreicherung einer bestimmten mRNA aus einer solchen Präparation, die nur zu etwa 10% mRNA enthält, auf die hohe Spezifität der Selektion hin; zweitens wurde häufig beobachtet, daß nach Anreicherung von mRNA Schwierigkeiten bei der Translation großer mRNA-Species auftraten.

Etwa 50-100 μ g isolierte Gesamt-RNA verschiedener Gewebe wurden jeweils in ein Experiment eingesetzt, was etwa 5-10 μ g angereicherter mRNA entsprach. Um eine hohe Konzentration der mRNA in der Hybridisierungslösung zu erzielen (etwa 1 μ g/ μ l), wurde das Volumen möglichst gehalten (100-200 μ l). In allen Hybridselektionsexperimenten wurden identische Bedingungen verwendet: 50% Formamid/ 5x SSC/ 37°C (78). Nach der Hybridisierung (24-36 Stunden) wurde die nicht gebundene RNA durch mehrmaliges Waschen mit 5x SSC und danach mit 50% Formamid / 5xSSC bei 37°C entfernt.

Abb. 14 : Hybridselektion von p550-spezifischer mRNA und In Vitro Translation (Kaninchen Reticulocyten Lysat)
Selektion aus RNA von A. Sterna/ 12 Tage

B. Sterna/ 19 "

C. Sehne / 16 "

Translation : K Ausgangs-RNA (1 μ g)

Ø ohne Zugabe exogener RNA

A Fraktion A = vor Elution

spezifisch gebundener mRNA

B Fraktion B = p550-spezifische mRNA

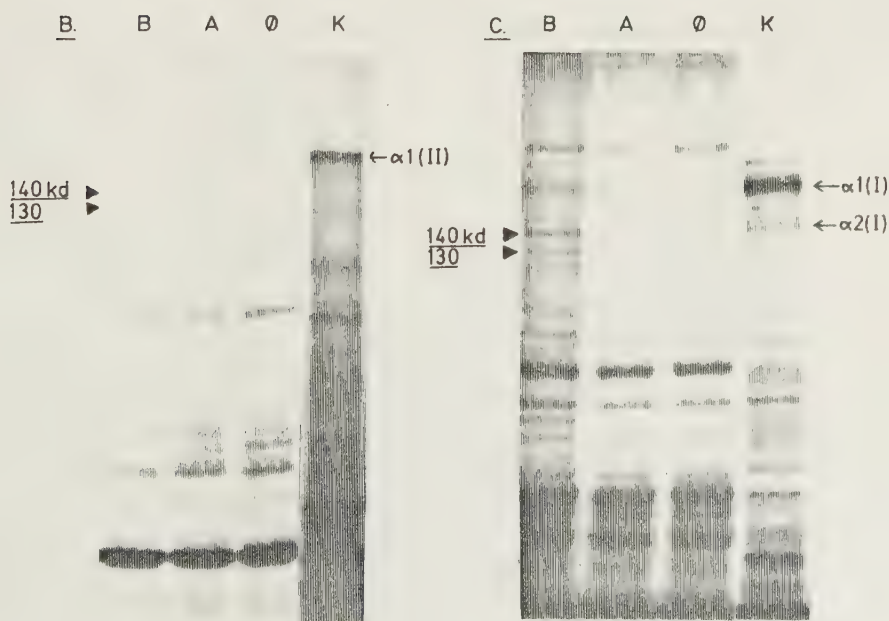
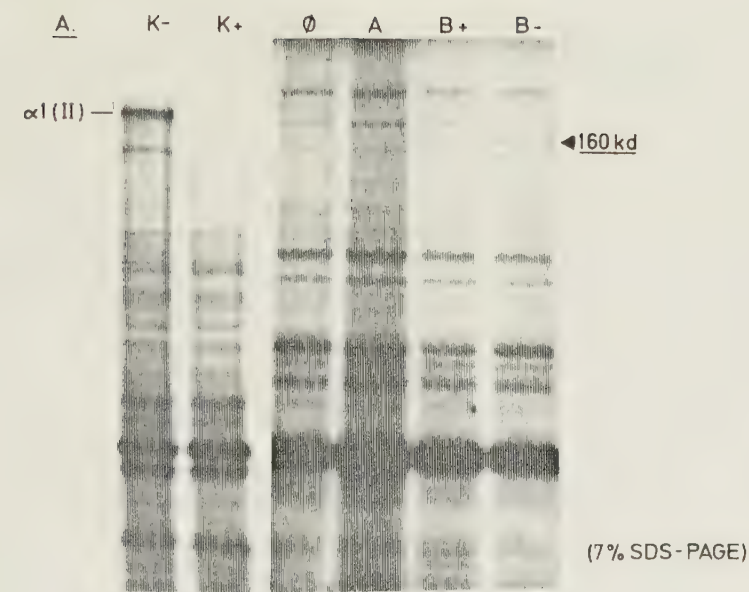
+/- mit/ohne Spaltung mit bakt.

Kollagenase

Gele : 7% SDS-Acrylamidgele

► : Spezifisch selektierte Proteine

Abb. 14A. B, C



Die Elution der gebundenen mRNA erfolgte in zwei Stufen. Mit 95% Formamid bei 37°C wurden alle lose gebundenen mRNAs abgelöst (Fraktion A). Erst danach wurde die spezifisch gebundene mRNA mit 95% Formamid bei 70°C freigesetzt (Fraktion B). Die RNA dieser beiden Fraktionen wurde durch Alkoholfällung konzentriert und dann in einem In Vitro Translations-system (aus Kaninchen Retikuloeyten) in Protein übersetzt. Die neu gebildeten Proteine der In Vitro Translation wurden durch den Einbau von ^{35}S -Methionin radioaktiv markiert und konnten nach Auftrennung in SDS-Acrylamidgelen durch Fluorographie sichtbar gemacht werden.

Das Ergebnis der Hybridselektion von Sterna-RNA (12 und 19 Tage alten Embryonen) und Sehnen-RNA (16 Tage) ist in Abb. 14 gezeigt. In allen Fällen wurde neben den Fraktionen A und B auch die zur Selektion verwendete RNA Präparation und eine Probe ohne Zugabe exogener mRNA translatiert. Die Größenbestimmung der Proteine erfolgte durch den Vergleich mit radioaktiven Markerproteinen. Sollten bei der Translation der Fraktion B Banden zu finden sein, die weder in der Fraktion A (lose gebundene RNA), noch in der Leerprobe (Background des Translationssystems) auftreten, so deutet dies auf die Translation einer spezifisch angereicherten mRNA hin.

Diese Kriterien erfüllt im Falle der Selektion von RNA aus Sterna/ 12 Tage nur ein Protein von etwa 160 kd (Abb. 12A) Bei Sterna/ 19 Tage (Abb. 12B) und Sehne/16 Tage (Abb. 12C) finden sich jedoch zwei Proteine von 140 und 130 kd, die in einem annähernd aquimolaren Verhältnis auftraten. In allen Fällen war es nicht möglich, die gefundenen Proteine im Translationsmuster der nicht selektierten RNA eindeutig zu identifizieren. Eventuell sind dafür die Bedingungen der Translation verantwortlich, da in Einzelversuchen mit geändertem Na^+/K^+ -Verhältnis eine bessere Translation beobachtet werden konnte. Insgesamt waren die Ausbeuten an translatierbarer mRNA gering, da bei der Selektion von 100 µg Gesamt-RNA die mRNA meist nur für 1-2 Translationen ausreichte.

In weiteren Versuchen sollte die Behandlung der in vitro translatierten Proteine mit bakterieller Kollagenase (aus Clostr. Histolyticum) zeigen, ob die entstehenden Proteine kollagene Sequenzen enthalten. Solche Sequenzen finden sich nicht nur bei Kollagenen, sondern konnten auch bei einigen minoren Matrixkomponenten nachgewiesen werden (25, 122,123). Für eine Spaltung mit diesem Enzym ist bereits die Sequenzfolge X-Pro-X-Gly-Pro ausreichend (80). Nur für das 160 kd Protein aus Sterna/12 Tage konnte eine Spaltung mit Kollagenase gezeigt werden (Abb. 12A), was auf die Existenz von zumindest kurzen kollagenen Sequenzen hinweist. Wieviele Fragmente bei der Spaltung entstehen, ist jedoch nicht zu erkennen. Ob auch die 140 kd und 130 kd Banden mit Kollagenase abbaubar sind, ist noch unklar, da die selektionierte mRNA nicht in ausreichender Menge zur Verfügung stand.

Tabelle 3 : Ergebnis der Hybridselektion p550-spezifischer mRNA und In Vitro Translation

| Gewebe | Entwicklungs-
stufe | Selektierte
Proteine | Kollagenase-
abbau |
|-----------|------------------------|-------------------------|-----------------------|
| Sternum | 12 Tage | 160 kd | + |
| | 19 " | 130/140 kd | n.b. |
| Kalvarien | 12 " | -? | |
| | 19 " | -? | |
| Sehne | 16 " | 130/140 kd | n.b. |

Die Ergebnisse der verschiedenen Experimente sind in der Tabelle 3 zusammengefaßt. Die Selektion von Kalvarien RNA (12 und 19 Tage) brachte kein Ergebnis. Ob in diesen Fällen keine translatierbare mRNA vorliegt oder ob nur technische Probleme für dieses Ergebnis verantwortlich sind, läßt sich nicht entscheiden.

Diese Proteine werden später mit bereits beschriebenen

Komponenten der verschiedenen Gewebe verglichen (s.G.1.). Welcher Zusammenhang zwischen den beiden mRNA-Species (4.3 kb, 6.5 kb) und der Synthese von drei Proteinen mit unterschiedlicher Länge (130, 140 und 160 kd) im In Vitro Translationssystem besteht, muß noch offen bleiben. Einige Aspekte werden später im Zusammenhang mit der Genstruktur und dem Expressionsmuster des p550 Gens diskutiert (s.G.5.).

2. Konstruktion gewebespezifischer Genbanken

2.1 Prinzip der Konstruktion von Genbanken

Ein Einblick in die Regulation der Expression des p550 Gens, seiner Transkription und des Processings der mRNA ist erst möglich, wenn die Struktur des zugehörigen Gens bekannt ist. Eine Erklärung der bisherigen Befunde wird erst dadurch möglich. Aus diesem Grund wurden Teile des p550-Gens aus einer Genbank isoliert.

Eine Genbank repräsentiert das gesamte Genom eines Organismus in Form einzelner, klonierter Fragmente und macht erst dadurch die Isolation bestimmter Bereiche des Genoms möglich. Zur Klonierung großer genomischer Fragmente verwendet man entweder Phagenvektoren, die von dem Phagen λ abgeleitet wurden (81) oder Cosmidvektoren (82). In den folgenden Versuchen wurde der 'Replacementvektor' λ Ch 4A (83) verwendet, der die Klonierung von Eco RI Fragmenten mit einer Länge von 8.2 - 22.2 kb erlaubt. Dieser Vektor erfüllt in der Kombination mit dem Empfängerstamm DP 50supF (84) die Richtlinien der biologischen Sicherheit (85).

Eine gleichmäßige Verteilung von genomischen Fragmenten eines bestimmten Längenbereichs läßt sich auf zwei Weisen erreichen: entweder durch partielle Spaltung mit häufig spaltenden Restriktionsenzymen (86) oder durch mechanisches Scheren hochmolekularer DNA. Zur Konstruktion der hier beschriebenen Genbanken wurde die Methode der Partialsplattung genomischer DNA durch Eco RI und die direkte Ligation der Fragmente in den Vektor verwendet (87). Versuche mit gescherter DNA wurden ebenfalls unternommen, führten jedoch nicht zur Konstruktion von repräsentativen Genbanken und werden deshalb nicht genauer beschrieben.

Das experimentelle Vorgehen zeigt Abb. 15. Der Vektor λ Ch 4A wird durch Eco RI in 4 Fragmente gespalten. Die beiden Vektorarme L (19 kb) und S (11 kb) werden von den beiden inneren Fragmenten (6 und 7 kb) abgetrennt und können mit genomischer DNA ligiert werden. Genomische DNA wird mit Eco RI partiell gespalten und Fragmente mit einer Länge von 15–22 kb werden abgetrennt. Die Verwendung von Fragmenten dieser Größenordnung verhindert, daß zwei Fragmente gemeinsam kloniert werden können. Es entstehen sehr lange konkatemere Moleküle durch die Ligation der cos-Enden untereinander und der Eco RI Enden von Insert-DNA mit den Vektorarmen. Die Selektion der rekombinanten Phagen erfolgt beim nächsten Schritt, dem Verpacken dieser DNA in Phagenköpfe, beim 'In Vitro Packaging'. Nur Moleküle mit einer Mindestlänge von 38 kb ergeben vitale Phagen, die nach der Transfektion von DP50 supF Plaques ausbilden können. Dieser Empfängerstamm supprimiert die Amber-Mutation des A-Gens des Vektors λ Ch4A.

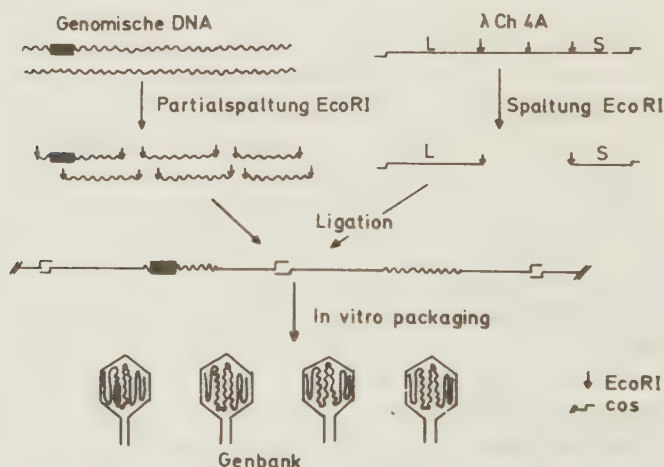


Abb. 15 : Prinzip der Konstruktion einer (Eco RI - Partial) Genbank in λ Ch 4A

Um eine statistisch repräsentative Genbank zu erhalten, ist eine Mindestanzahl von rekombinanten Phagen nötig, die sich aus der Genomgröße und der mittleren Insertlänge ableiten läßt. Die Genomgröße des Huhns beträgt für den diploiden Chromosomensatz etwa 2.4×10^9 bp (88). Damit läßt sich die Anzahl der rekombinanten Phagen berechnen, die nötig ist, damit jedes mögliche Fragment mindestens einmal in der Genbank vertreten ist.

Zur Berechnung wurde die Formel von Clarke & Carbon (89) verwendet :

$$N = \ln (1-P) / \ln (1-f)$$

wobei : N = Zahl der nötigen Rekombinanten

P = Wahrscheinlichkeit des Auftretens

f = Relation Insert/Genom

Damit jedes Fragment (mittlere Länge etwa 1.8×10^4 bp) mit großer Sicherheit ($P = 0.99$) in der Genbank enthalten ist, ergibt sich folgende Berechnung :

$$f = 1.8 \times 10^4 / 2.4 \times 10^9 = 7.5 \times 10^{-6}$$

$$N = \ln (1-0.99) / \ln (1-7.5 \times 10^{-6}) = 610\ 000$$

Eine vollständige Genbank muß damit mindestens 6×10^5 Phagen enthalten.

Gewebespezifische Genbanken

Bei den in der Literatur beschriebenen Genbanken des Hühnergenoms wurde meist von Erythrocyten-DNA ausgegangen (90). Diese Genbanken liegen nicht mehr in der Ausgangsform vor, sondern wurden bereits mehrfach amplifiziert, um höhere Titer zu erhalten. Dabei ist es möglich, daß durch unterschiedliche Wachstumsraten einzelner Phagen keine gleich-

mäßige Verteilung der Fragmente in der Genbank mehr vorhanden ist. Aus diesem Grund wurden diese Genbanken nicht verwendet.

Eigene, gewebespezifische Genbanken wurden noch aus einem weiteren Grund hergestellt. Das p550-Gen wird nur in einigen speziellen Bindegeweben exprimiert (s.C.1.4.). Durch die Klonierung der DNA von Chondrocyten konnte man annehmen, das funktionsfähige authentische Gen zu isolieren. Es deuteten bisher zwar noch keine Ergebnisse auf Unterschiede in der Genstruktur verschiedener Gewebe hin. Um solche denkbaren Unterschiede jedoch erfassen zu können, wurde DNA von *Sterna* (p550-Gen exprimierend) und von bereits dedifferenzierten Chondrocyten (nicht exprimierend) zur Konstruktion von Genbanken verwendet.

2.2 Isolation genomischer DNA und Partialspaltung mit EcoRI

Sternaspezifische DNA wurde aus freigesetzten Chondrocyten der *Sterna* von 17 Tage bebruteten Hühnerembryonen gewonnen. Das umgebende Perichondrium wurde durch Behandlung mit Kollagenase und Trypsin abgebaut (s.G.2.2.) und durch weitere Inkubation mit Kollagenase und Trypsin ließen sich aus 400 *Sterna* etwa 10^9 Chondrocyten freisetzen. Es war nicht mehr nötig, die Zellzahl durch kurzzeitige Kultur zu erhöhen, da eine ausreichende Menge an DNA gewonnen wurde.

Durch Zugabe eines Detergens (Sarkosyl) und EDTA ließen sich die Zellen lysieren; die Proteine wurden durch Proteinase K und die RNA durch RNase A abgebaut. Danach wurde die DNA in einem Dichtegradienten weiter gereinigt. Mit dieser Methode konnten etwa 3 mg hochmolekulare DNA isoliert werden, deren Länge etwa 100 kb betrug.

Ein Teil der isolierten Chondrocyten wurde weiter kultiviert (s.G.2.1.). Nach etwa sechs Wochen hatte die überwiegende Anzahl der Zellen auf die Synthese von Typ I Kollagen umgeschaltet und damit den differenzierten Zustand des Knorpels verloren. Die Umschaltung war durch Immunfluoreszenz

mit typspezifischen Antikörpern der Kollagene verfolgt worden. Die DNA dieser Zellen wurde ebenfalls auf die vorher beschriebene Weise freigesetzt und weiter gereinigt.

Partialsplaltung mit Eco RI

Um eine möglichst repräsentative Verteilung aller möglichen Partialfragmente im Größenbereich von 15-22 kb zu erhalten, wurde die DNA unter verschiedenen Bedingungen mit Eco RI gespalten. In Abb. 16 ist ein solcher Abbau gezeigt. Eine deutlich unterschiedliche Verteilung der Fragmentlängen ist bei Änderung von Enzymmenge und Dauer der Inkubation zu erkennen. Die Bedingungen dieses Vorversuchs wurden auch für die präparativen Partialsplaltungen verwendet.

Jeweils 100 µg der genomischen DNA, isoliert aus *Sterna*, wurden analog dem Vorversuch (Abb. 16), mit 0.08, 0.16, 0.32 und 0.48 U Eco RI/ µg DNA für 30 Min gespalten. Die Reaktion wurde durch Phenolextraktion gestoppt und die DNA durch eine Alkoholfällung konzentriert. Die verschiedenen Ansätze wurden gemischt, in einem Sucrosegradienten (10-40%) aufgetrennt und fraktioniert. Die Länge der DNA-Fragmente in den einzelnen Fraktionen wurde durch Gelelektrophorese bestimmt (Abb. 17). Alle Fraktionen, die Fragmente mit einer Länge von 15-22 kb enthielten, wurden vereinigt, dialysiert und alkoholgefällt. Die so gereinigten Partialfragmente sind in Abb. 18 gezeigt.

Auf die gleiche Weise wurden Partialfragmente, ausgehend von der DNA der dedifferenzierten Chondrocyten, hergestellt.

Bei einer statistischen Verteilung aller Basen innerhalb einer Sequenz sollte eine Eco RI Stelle (GAATTC) jeweils nach etwa 4,1 kb (1 von 4⁶ Möglichkeiten) auftreten. Sind zwei Eco RI Stellen jedoch weiter als 22 kb voneinander entfernt kann ein solches Fragment nicht kloniert werden und ist deshalb niemals in der Genbank vertreten. Bei einer Häufung von Eco RI Stellen dagegen, finden sich unter Umständen keine überlappenden Fragmente im verwendeten Größenbereich von 15-22 kb. Durch die Klonierung von Eco RI-Partialfrag-

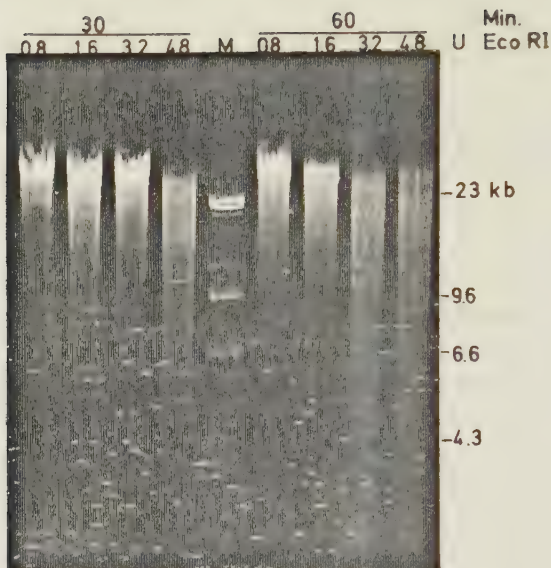


Abb. 16 : Partialsplaltung genomischer DNA mit Eco RI :
 Sterna-DNA ($10\mu\text{g}/50\mu\text{l}$) wurde mit 0.8-4.8 U Eco RI
 für 30 bzw. 60 Min. bei 37°C abgebaut.
 (0.5% Agarosegel) M : λ / Hind III

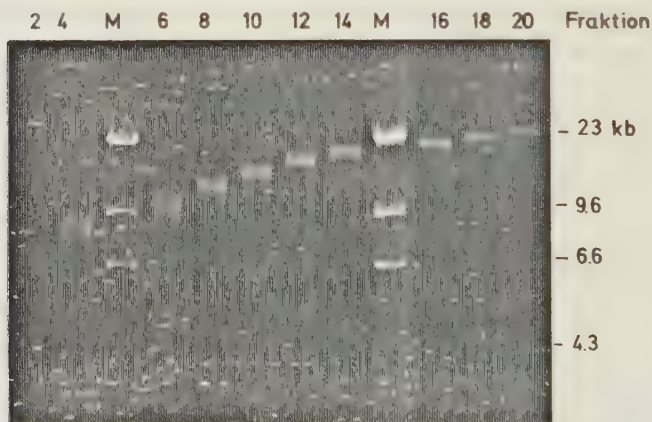


Abb. 17 : Fraktionierung Eco RI partial-gesplattener Sterna-DNA durch Sucrosegradientenzentrifugation :
 Partialgesplattene Sterna DNA wurde in einem 10-40%
 Sucrosegradienten zentrifugiert, fraktioniert und
 Aliquots in einem 0.5% Agarosegel aufgetrennt.
 M : λ / Hind III

menten ergibt sich damit eine Beschränkung in der gleichmäßigen Verteilung von Fragmenten in den Genbanken.

2.3 Klonierung in λ Ch 4A

Der Vektor λ Ch 4A wurde durch Eco RI in 4 Fragmente gespalten. Die beiden 'inneren' Eco RI Fragmente (6 und 7 kb lang) von λ Ch 4A enthalten keine für das Wachstum des Phagen notwendigen Gene und können deshalb durch genomische Fragmente ersetzt werden. Die beiden langen Vektorarme, L (19 kb) und S (11 kb), die zur Ligation nötig sind, wurden nach der Spaltung der Phagen-DNA mit Eco RI durch Zentrifugation in einem Sucrosegradienten von den inneren Fragmenten (6 und 7 kb) abgetrennt. Die verwendete Präparation von Vektorarmen ist in der Abb. 18, nach elektrophoretischer Auftrennung in einem Agarosegel, gezeigt.

Zur Ligation mit dieser Präparation von Phagenarmen wurden die isolierten Partialfragmente (Abb. 18) mit einer Länge von 15-22 kb verwendet. Zur Ligation wurden Insert-DNA und Vektorarme in einem molaren Verhältnis von 1,4 : 1 eingesetzt, wobei die Konzentration der Vektor-DNA im Ansatz etwa 170 μ g/ml betrug. Auf Grund theoretischer Vorhersagen ist ein molares Verhältnis von etwa 1.4 - 1.7 nötig, um optimale Ausbeuten zu erzielen (91).

Vor Zugabe der Ligase wurde der gesamte Ansatz bei 42°C inkubiert, um das 'Annealing' der cos-Enden der Vektorarme zu begünstigen (86), und damit die Entstehung langer, konkatemerer Moleküle zu erleichtern, die für ein optimales Verpacken der DNA nötig sind (92). Erst nach diesem Schritt wurde die Ligation gestartet (s.G.2.6.), die zur Entstehung sehr langer Moleküle führte, was durch Gelelektrophorese verfolgt werden konnte.

Zusätzlich wurden auch Vektorarme alleine unter den gleichen Bedingungen ligiert. Die Anzahl der nach dem In Vitro Packaging auftretenden Phagen zeigte den Anteil von Phagen in der Genbank, der durch den Einbau noch vorhandener

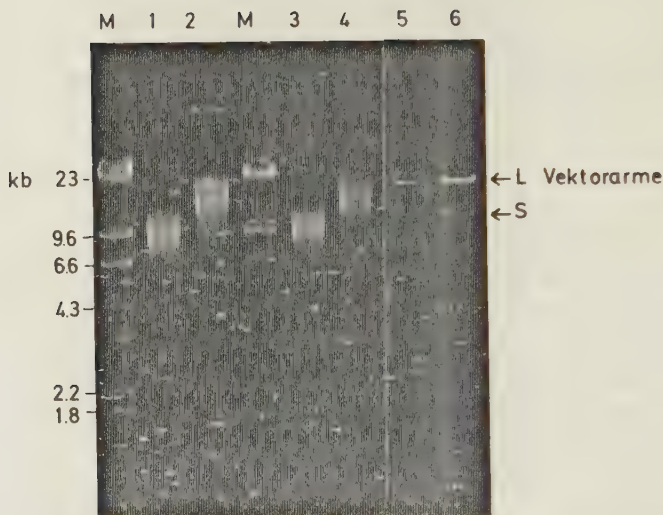


Abb. 18 : Vektorarmpräparation und Eco RI partial-gespaltene genomische DNA

- 1 : Sterna-DNA/ 8-15 kb
- 2 : " / 15-22 kb
- 3 : Dediff. Chondrocyten-DNA/ 8-15 kb
- 4 : " / 15-22 kb
- 5/6 : Vektorarmpräparationen (λ Ch4A)
- M : λ / Hind III (0.5% Agarosegel)

innerer Fragmente entstanden ist (s.C.2.4.).

In der Tabelle 4 (s.C.2.4.) sind die Bedingungen der Ligation und die Ausbeuten des In Vitro-Packaging zusammengefaßt.

2.4 In Vitro Packaging der Genbanken

Um intakte, infektiöse Phagen zu erhalten, muß die ligierte DNA in Phagenköpfe verpackt werden (92), die nach der Transfektion eines geeigneten Empfängerstamms zur Ausbildung von Plaques führen. Ein Extrakt, der dazu in der Lage ist, läßt sich durch Induktion der beiden λ -lysogenen Stämme BHB 2688 und BHB 2690 (92) herstellen. Beide Stämme enthalten Amber-Mutationen in jeweils einem der Kapsidpro-

teine. Dadurch bleibt nach der Induktion (Temperaturschock 42°C) von BHB 2688 die Bildung von Phagenpartikeln auf einer Zwischenstufe stehen, da das Kapsidprotein D nicht gebildet werden kann. In BHB 2690 entstehen dagegen keine Phagenköpfe, da hier das Hauptprotein des Phagenkopfes (E) nicht gebildet wird. Da beide Stämme aufgrund einer Mutation des S-Gens nicht lysieren können, häufen sich die phagenspezifischen Produkte in den Zellen an. Die Mischung der Lysate beider Stämme rekonstituiert das gesamte System und es können intakte Phagenköpfe gebildet und DNA darin verpackt werden.

Der 'In Vitro Packaging Extrakt' besteht aus einer 1:1 Mischung von Kulturen der beiden induzierten, aber noch nicht lysierten Stämme. Die Lyse der Bakterien, die Freisetzung der Phagenkomponenten erfolgt erst durch Tauen des tiefgefrorenen Extraktes direkt vor der Zugabe der zu verpackenden DNA. Um den Einbau endogener λ -DNA und auch eine eventuelle Rekombination mit bakterieller DNA zu verhindern, wird die Mischung der induzierten Stämme ausgiebig mit UV bestrahlt. Die zusätzlich vorhandene rec A^{-} Mutation in beiden Stämmen verhindert mögliche Rekombinationsereignisse.

Die Herstellung des Packaging Extrakts (s.G.2.6.) erfolgte in enger Anlehnung an die Methode von Hohn (92). Nach dem Wachstum und der Induktion (bei 42°C) der beiden Stämme BHB 2688 und BHB 2690 wurden die Kulturen gemischt und mit UV (260 nm) bestrahlt. Die Zeitdauer und Intensität der Bestrahlung, die nötig ist, um den Titer einer Suspension vitaler $\lambda\text{Ch 4A}$ Phagen um einen Faktor von 10^{-8} zu senken, wurde in Vorversuchen empirisch bestimmt. Diese Bedingungen entsprechen den Richtlinien der biologischen Sicherheit (85).

Die Verpackungseffizienz so hergestellter Packaging Extrakte betrug $0.5 - 2 \times 10^8$ pfu/ μg DNA ($\lambda\text{Ch 4A}$). Die Effizienz des Packaging konkatemerer Moleküle liegt noch höher, da Einzelmoleküle kein optimales Substrat darstellen.

Packaging der Genbanken

Aliquote Anteile der Ligationsansätze (5 μ l) wurden nach dem Auftauen der Extrakte, und damit der Freisetzung der für das Verpacken nötigen Komponenten mit der sehr viskosen Lösung vorsichtig verrührt und bei 37°C inkubiert. Um eine möglichst vollständige Verpackung zu erreichen, wurde nach 45 Minuten nochmals frischer Extrakt zugegeben. Zusätzlich wurde der Ansatz mit DNase I behandelt, da dadurch eine noch höhere Effizienz erreicht werden kann (92).

Tabelle 4 : Ligation und Packaging der Genbanken

| <u>Ligationen</u> : | Kontrolle | 'Sterna' | 'Dediff.' |
|---------------------------------|-------------------|--------------------|-------------------|
| Volumen Ligation (μ l) | 5 | 30 | 30 |
| μ g λ Ch 4A Arme | 0.85 | 5.1 | 5.1 |
| μ g/ml λ Ch 4A Arme | 170 | 170 | 170 |
| μ g Insert DNA (15-22 kb) | - | 4.2 | 4.2 |
| μ g/ml Insert DNA | - | 140 | 140 |
| molaes Verhältnis | | | |
| Insert/Vektor | 0 | 1.4 | 1.4 |
| <u>Packaging</u> : | | | |
| Gesamt pfu | 3.2×10^4 | 5.25×10^6 | 5.1×10^6 |
| pfu/ μ g Vektor | 2.9×10^4 | 1.0×10^6 | 1.0×10^6 |
| Anteil Kontrolle | 3% | 100% | |
| <u>Reinigung</u> : | | | |
| Gesamt pfu (nach | | | |
| Gradient) | - | 1.8×10^6 | 1.5×10^6 |
| Genomäquivalente | | 3 | 2.5 |
| <u>Amplifikation</u> : | | | |
| Ausgangstitel (pfu) | | 0.8×10^6 | |
| Endtitel (pfu) | | 3.8×10^8 | |
| x Amplifikation | | 480x | |

Nach der Zugabe von Phagenpuffer und Chloroform konnte der Phagentiter im Überstand (nach der Zentrifugation des Bakterienpellets) durch Transfektion von DP 50 supF und Plattierung auf einen Bakterienrasen bestimmt werden (s.G.2.7.). Die Tabelle 4 zeigt eine Zusammenfassung von Ligation und Packaging der Genbanken.

In einem CsCl Stufengradienten wurden die Phagen nach dem Packaging weiter gereinigt, wobei ein Teil der Phagen verloren ging. Die verbliebene Anzahl von Phagen war jedoch ausreichend für eine repräsentative Genbank, für die etwa 6×10^5 Phagen nötig sind (s.C.2.1.).

Aus Ligation und Packaging der Kontrolle ließ sich schließen, daß etwa 3% der entstandenen Phagen keine echten Rekombinanten darstellten, sondern unvollständig abgetrennte innere Fragmente des Vektors eingebaut enthielten. Der Background der Genbanken wurde später genauer untersucht (s.C.2.5.).

Amplifikation der Genbank

Ein Teil der Phagen der 'Sterna' Genbank wurde amplifiziert. Dazu wurden 8×10^5 pfu der primären 'Sterna' Genbank in einer Dichte ausplattiert, die noch keine konfluente Lyse ergab, jedem Phagen also die gleichen Wachstumsbedingungen bot. Durch Diffusion ließen sich die Phagen aus der Agarschicht freisetzen und konnten durch einen anschließenden Dichtegradienten gereinigt werden. Es wurde dabei eine etwa 500-fache Amplifikation des Ausgangstiters erzielt (s.Tabelle 4).

2.5 Background der Genbanken

Der Einbau von 'inneren' Fragmenten läßt sich genauer untersuchen, wenn die Phagen auf das Vorhandensein des lac^+ und des bio^+ -Gens getestet werden, die auf dem 6 kb bzw. dem 7 kb Eco RI-Fragment von $\lambda\text{Ch 4A}$ lokalisiert sind. Zu diesem Zweck wurde DP50 supF mit den Phagen der Genbanken infiziert und auf Minimalagarplatten/-Biotin/+X-Gal kultiviert. Phagen, die das 6 kb Fragment eingebaut enthalten, erkennt man deshalb an der Blaufärbung der Plaques. Das Vorhandensein des bio^+ -Gens des 7 kb Fragments resultiert in einer Zone dichteren Bakterienwachstums in der Umgebung der Plaques. Enthält ein Phage beide Fragmente, findet man blaue Plaques, umgeben von einem Wachstumshof. Das Ergebnis des Tests von 500-1000 Plaques ist in Tabelle 5 zusammengefaßt.

Tabelle 5 : Background der Genbanken

| Genbank | 'Sterna' | 'Dediff.' |
|--|----------|-----------|
| 6 kb/ lac^+ | 1.5% | 1% |
| 7 kb/ bio^+ | 3% | 4% |
| lac^+ / bio^+ | 0.5% | 1% |
| Gesamt | 4% | 4% |
| (lac^+ oder bio^+ oder
$\text{lac}^+/\text{bio}^+$) | | |

Etwa 4% aller Phagen hatten zumindest eines der beiden inneren Fragmente eingebaut. Da bei der Ligation und dem Packaging der 'Kontrolle' (Tab. 4; s.C.2.4.) ein Background von etwa 3% gefunden wurde, enthielten damit etwa 1% aller Phagen neben einem genomischen Fragment noch eines der beiden inneren Fragmente. Dieser Anteil an 'irregularen' Klonen beeinflußt die Repräsentativität der Genbanken noch nicht. Die Genbanken wurden deshalb in den folgenden Versuchen weiter verwendet.

3. Isolierung und Restriktionskartierung genomischer Klone des p550 Gens

3.1 Prinzip der Isolierung genomischer Klone

Rekombinante Klone einer Genbank, die Teile eines bestimmten Gens eingebaut enthalten, lassen sich an der Hybridisierung mit einer spezifischen (cDNA-) Probe erkennen. Die Phagen werden dazu in Form individueller Plaques kultiviert und danach auf Nitrocellulosefilter transferiert. Durch die Diffusion eines Teils der Phagen eines Plaques aus der Agarschicht der Platten auf die Filter entstehen Replikas der Ausgangsplatten. Nach der Fixierung der Phagen auf den Filtern durch Trocknung muß die DNA aus den Phagenköpfen freigesetzt und denaturiert werden, um sie einer Hybridisierung zugänglich zu machen. Durch einen Denaturierungsschritt mit 200 mM NaOH werden sowohl die Phagenköpfe zerstört als auch die DNA in die Einzelstränge zerlegt. Nach dem Neutralisieren und Waschen kann die nun einzelsträngige DNA an der Nitrocellulose durch Backen fixiert werden. Die Filter können nun mit der radioaktiv markierten Probe hybridisiert werden.

Für eine sichere Identifizierung von positiven Klonen ist es notwendig, Duplikate der Filterreplikas herzustellen. Nur wenn auf den parallel behandelten Filtern der gleichen Platte an identischen Stellen hybridisierende Plaques zu finden sind, kann von einem positiven Klon gesprochen werden.

Für die praktische Durchführung des Screenings einer Genbank treten zwei entgegengesetzte Probleme auf: erstens möchte man nur eine geringe Anzahl von Phagen testen, um den Aufwand an Platten und Filtern möglichst gering zu halten; zweitens steigen aber mit der Anzahl der getesteten Phagen

die Aussichten auf die Isolation eines Klons.

Die Anzahl der Phagen, die nötig ist, um erfolgreich in den konstruierten Genbanken des Hühnergenoms nach dem p550-Gen zu suchen, läßt sich mit Hilfe der Formel von Clarke & Carbon (89) berechnen.

$$N = \ln(1-P) / \ln(1-f) \quad (\text{s.C.2.1.})$$

Damit lassen sich nach der Umformung der Formel die Wahrscheinlichkeiten P berechnen, mit der ein positiver Klon in einer Stichprobe N der Genbank zu finden ist:

| $P = 1 - e^{N \times \ln(1-f)}$ | für | N (pfu) | P |
|---------------------------------|-----|-----------------|------|
| | | 1×10^5 | 0.53 |
| | | 2×10^5 | 0.78 |
| | | 3×10^5 | 0.90 |
| | | 6×10^5 | 0.99 |

Mit der Größe der Stichprobe, also der Anzahl der getesteten Phagen, nimmt die Wahrscheinlichkeit schnell zu, einen positiven Klon zu entdecken. Etwa 200.000 Phagen sollten für das Screening dieser Genbank zumindest verwendet werden, um sowohl eine gute Chance (78%) zu haben, einen Klon zu entdecken, als auch den Aufwand möglichst gering zu halten.

3.2 Screening der 'Sterna' Genbank mit dem cDNA Klon p550

Für das Screening der 'Sterna' Genbank mit dem cDNA Klon p550 wurden auf Grund der vorherigen Überlegungen 200.000 pfu verwendet. Diese Anzahl von Phagen ließ sich auf 20 Platten ($\phi = 15$ cm) kultivieren (10 000 pfu/Platte). Von jeder Platte wurden zwei identische Filterreplikas hergestellt (insgesamt 40 Filter) und parallel weiter behandelt (s.G.2.7.).

Für die Hybridisierung der Replikafilter wurde das Insert von p550 isoliert und durch Nicktranslation, unter Verwendung von $\alpha^{32}\text{P}$ -dCTP radioaktiv markiert. Es wurde eine spezifische Radioaktivität von etwa 1.6×10^8 cpm/ μg DNA wurde dabei erreicht. Jeweils 10 Filter wurden zusammen mit 1×10^6 cpm/ml Hybridisierungslosung nach der 'Sarkosylmethode' hybridisiert und danach gewaschen (s.G.2.5.).

In Abb. 19 ist die Autoradiographie zweier parallel hergestellter Filter gezeigt. Auf beiden Filtern ist an identischen Stellen ein deutliches Signal zu erkennen. An der zugehörigen Stelle der Ausgangsplatten befindet sich somit ein rekombinanter Phage, der Sequenzen des cDNA Klonen p550 enthält: in diesem Fall $\lambda\text{St 550-3}$. Die Markierung von Platten und Filtern erlaubt zwar eine grobe Lokalisierung des positiven Klonen, aber aufgrund der Dichte der Plaques kann jedoch kein einzelner Plaque dem Signal zugeordnet werden.

Insgesamt wurden 5 positive Klone gefunden, die mit $\lambda\text{St 550-1 bis -5}$ bezeichnet wurden.

Reinigung der positiven Klone

Um einen positiven Klon zu reinigen, wurden etwa 100 Plaques in der Umgebung des gefundenen Signals mit sterilen Zahnstochern gepickt und auf einen Rasen von DP50 supF geimpft. Nachdem sich Plaques entwickelt hatten, wurden wieder Filterreplikas hergestellt und mit p550 hybridisiert. In diesem 1. Reinigungsschritt zeigen nur einige der Plaques eine positive Reaktion (Abb. 19). Die Isolation dieser Phagen ergab noch keine einheitliche Phagenpopulation.

Die Phagen eines positiven Plaques wurden deshalb isoliert, verdünnt und in einer Dichte auf einen Rasen von DP50 supF ausgestrichen, die zur Entstehung individueller Plaques führte. Einzelne Plaques wurden wieder gepickt, kultiviert, auf Filter transferiert und hybridisiert. Bei diesem 2. Reinigungsschritt zeigte nur noch ein kleiner Teil der Plaques keine Hybridisierung mit p550.

Diese Schritte wurden solange wiederholt, bis alle

Abb. 19 : Identifizierung und Reinigung genomischer Klone des p550-Gens in der 'Sterna-Genbank' (jeweils parallele Filter)

A. Screening der 'Sterna-Genbank' (10 000 pfu/Filter) durch Hybridisierung mit nick-translatiertem Insert von p550
Positives Signal : λ St 550-3

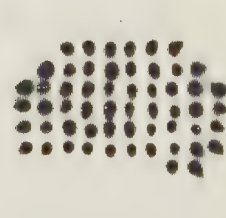
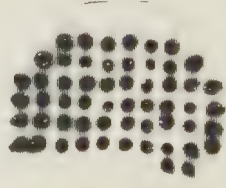
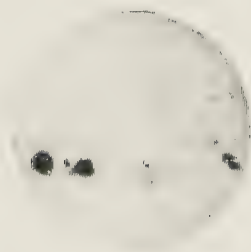


B. Schrittweise Reinigung des Klon λ St 550-3

1.

2.

3.



Plaques eine positive Reaktion zeigten. Dies war in den meisten Fällen nach dem 3. Reinigungsschritt der Fall. Erst nachdem 100% der Phagen eine Hybridisierung mit p550 zeigten, wurde von einem einheitlichen Klon gesprochen werden (s. Abb. 19B).

Mit dieser Methode wurden alle 5 positiven Klone des Screening mit p550 bis zur Homogenität gereinigt.

3.3 Restriktionskartierung der Klone λ St 550-1 bis -5

Die Restriktionskartierung der isolierten Klone erlaubt eine Aussage darüber, ob alle Klone gleiche genomische Fragmente enthalten, oder ob überlappende Genbereiche repräsentiert sind. In einem ersten Vorversuch wurden deshalb kleine Mengen der Phagen-DNA isoliert und mit verschiedenen Restriktionsenzymen gespalten.

Zur Präparation kleiner Mengen der Phagen-DNA wurden die Phagen in einer Dichte ausplattiert, die eine annähernd konfluente Lyse ergab (etwa $1-2 \times 10^4$ / 10 cm Platte). Bei der Überschichtung mit Phagenpuffer wanderte ein Großteil der Phagen in den Puffer, und konnte daraus mit PEG gefällt werden. Bakterielle DNA und RNA wurde vorher durch Inkubation mit DNase I und RNase A abgebaut. Aus dem so gereinigten Phagenpellet wurde die DNA durch Phenolextraktion freigesetzt und durch Alkoholfällung konzentriert.

Die isolierte DNA wurde zuerst mit Eco RI gespalten. Da Eco RI Fragmente in den Vektor eingebaut wurden, sollten nach der Spaltung die Eco RI Fragmente des klonierten Inserts neben den Vektorarmen (11 und 19 kb) zu erkennen sein. Das Ergebnis einer solchen Spaltung ist in der Abb. 20 gezeigt. Die Längen der sichtbaren Fragmente sind in der Tabelle 6 angegeben.

Bei der Spaltung mit Eco RI fanden sich neben den beiden Vektorarmen ausschließlich Fragmente, die dem klonierten genomischen Insert entsprachen. Die Klone λ St 550-1,-2 und -4 zeigten ein identisches Spaltungsmuster.

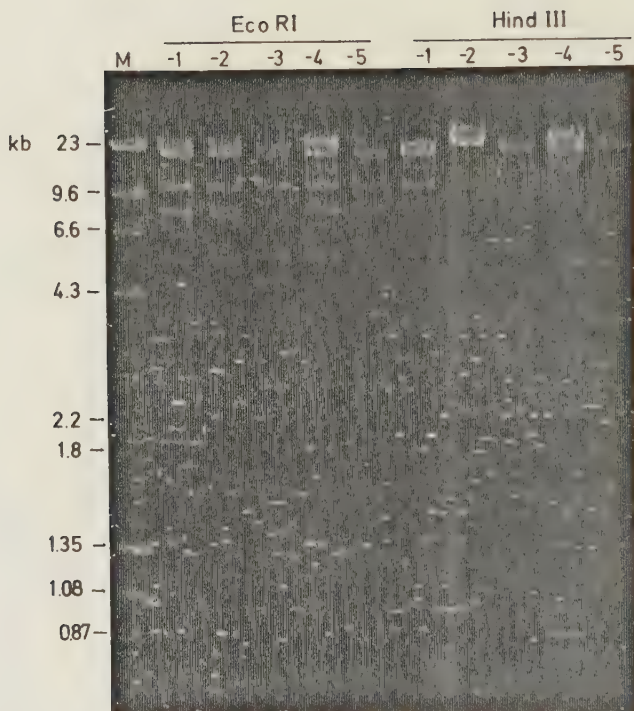


Abb. 20 : Spaltungen der genomischen Klone λ St 550-1 bis -5
mit Eco RI und Hind III
M : λ / Hind III + ϕ X/ Hae III
(0.6% Agarosegel)

Tabelle 6 : Längen der Eco RI Fragmente der Klone λ St 550-1,
-2, -3, -4 und -5

| Klon λ St 550- | -1 | -2 | -3 | -4 | -5 | |
|------------------------|-----|-----|-----|-----|-----|----|
| Vektorarm (L) | 19 | 19 | 19 | 19 | 19 | kb |
| (S) | 12 | 12 | 12 | 12 | 12 | |
| Insertfragmente | 8.1 | 8.1 | | 8.1 | 8.1 | |
| | 5.3 | 5.3 | 5.3 | 5.3 | 5.3 | |
| | | | 4.5 | | | |
| | | | 2.6 | | 2.6 | |
| | | | 1.3 | | | |
| | 0.7 | 0.7 | | 0.7 | 0.7 | |

Der Klon λ St 550-5 wies noch ein zusätzliches, 2.6 kb langes Eco RI Fragment auf. Von diesen vier ähnlichen Klonen wurde deshalb λ St 550-5 für eine detailliertere Restriktionsanalyse ausgewählt, da er offensichtlich den größten genomischen Bereich repräsentiert.

Der Klon λ St 550-3 zeigte im Vergleich dazu ein unterschiedliches Spaltmuster. Das 5.3 kb Fragment differierte geringfügig in der Länge von dem der anderen Klone. Mit dem Klon λ St 550-5 gemeinsam hatte es nur das 2.6 kb Fragment. Der Klon überlappte demnach in diesem Bereich mit λ St 550-5 und wurde deshalb ebenfalls für eine genaue Restriktionskartierung ausgewählt.

Daneben wurden die Klone mit Hind III gespalten, um die Orientierung der Vektorarme bezüglich des Inserts erschließen zu können. Der Vektor weist nur im S-Fragment eine Hind III-Stelle auf (93). Bei der Spaltung entsteht deshalb ein 5.7 kb langes Fragment des Vektoranteils, das bei allen Spaltungen sichtbar ist. Der deutliche Unterschied in den Spaltmustern von λ St 550-2 verglichen mit λ St 550-1 und -4 legt den Schluß nahe, daß in diesen beiden Fällen das Insert in unterschiedlicher Orientierung eingebaut wurde. Die Lage und Orientierung dieser Klone ist im Zusammenhang mit der Restriktionskarte der gesamten klonierten Region genau angegeben (s.C.3.4.; Abb. 24).

3.4 Restriktionskartierung der Klone λ St 550-3 und -5

Die Spaltung der Phagen-DNA mit verschiedenen Restriktionsenzymen erlaubt die Aufstellung von detaillierten Restriktionskarten. Zur groben Kartierung wurden Restriktionsenzyme verwendet, die nur selten spalten und deshalb wenige Fragmente erzeugen. Die Abb. 21A zeigt ein Beispiel solcher Spaltungen. Der Vergleich der Spaltungen verschiedener Kombinationen von Enzymen erlaubte die Lokalisierung der einzelnen Fragmente. Randfragmente, die Teile des Vektors

und des Inserts enthalten, ließen sich aus der bekannten Restriktionskarte von λ Ch 4A (93) ableiten.

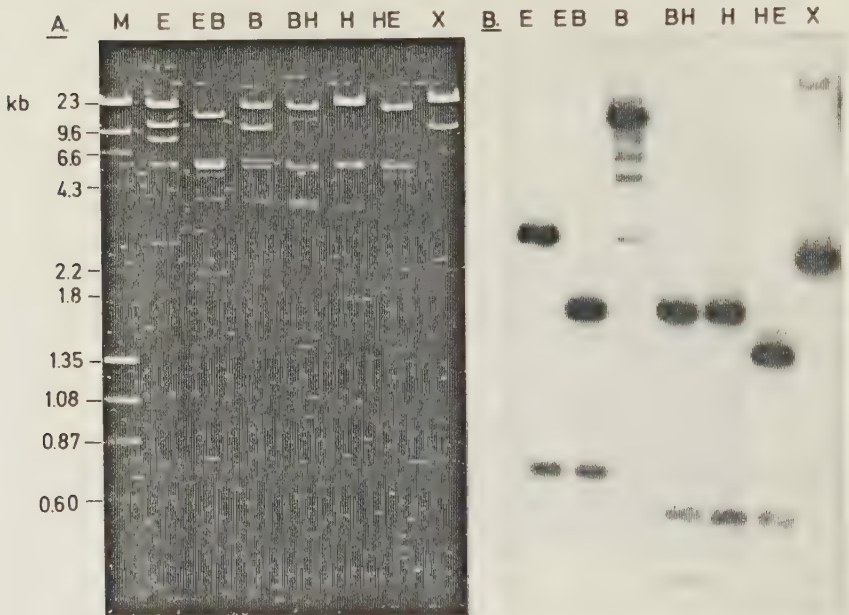


Abb. 21 : Restriktionskartierung und Southernblotting von λ St 550-5

A. Spaltung von λ St 550-5 mit verschiedenen Restriktionsenzymen (0.8% Agarosegel)

B. Southerntransfer und Hybridisierung mit dem markierten Insert von p550 (Autoradiographie)

E : Eco RI; B : Bam HI; H : Hind III; X : Xba I;
M : λ / Hind III + ϕ X/ Hae III

Die Frage nach der Lage der p550-Sequenz ließ sich durch Southernblotting (94) und Hybridisierung mit p550 beantworten. Nach der Auftrennung der Restriktionsfragmente von Spaltungen der DNA von λ St 550-3 und -5 in einem Agarosegel wurden die Fragmente denaturiert, auf eine Transfermembran transferiert und dort durch Backen der Filter (80°C) fixiert (s.G.2.3.). Diese Filter mit dem identischen Abbild der

Fragmenttrennung des Gels wurden mit dem, durch Nicktranslation markierten Insert des Klon p550 hybridisiert.

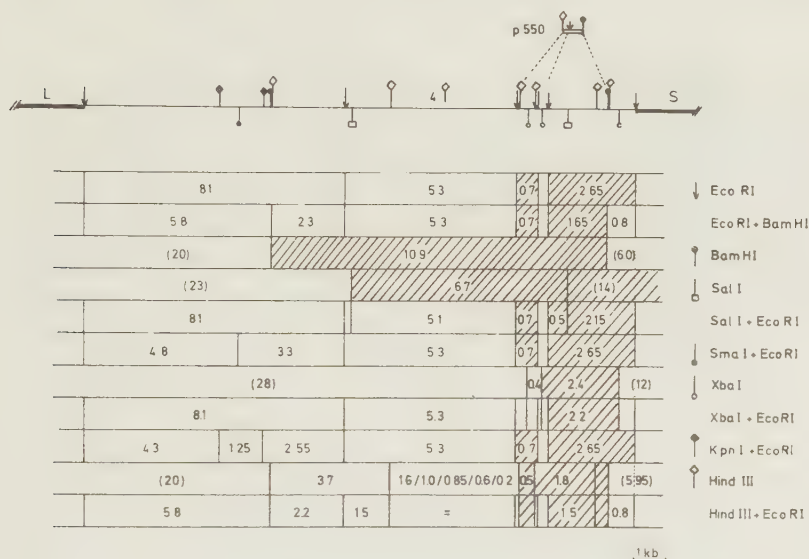


Abb. 22 : Restriktionskarte von λ St 550-5

L,S : Vektorarme von λ Ch 4A

(N) : Fragmente, die Teile des Vektors enthalten

/// : Hybridisierung mit p550 (keine Angaben bei Fragmenten unter 400 bp Länge)

--- : Zuordnung Restriktionsstellen p550 - Genom

Die Abb. 21 zeigt die Autoradiographie eines hybridisierten Filters neben dem zugehörigen, mit Ethidiumbromid gefärbten Agarosegel mit Spaltungen von λ St 550-5 vor dem Transfer. Fragmente, die p550-Sequenzen enthalten, ließen sich durch den Vergleich von Gel und Autoradiographie erkennen. Die Längen der Restriktionsfragmente, deren Anordnung und Hybridisierung mit p550 ist in der Abb. 22 zusammengefasst.

Für beide Klone ließen sich auf diese Weise eindeutige Restriktionskarten (Abb. 22 und 23). ableiten. Nur im Falle des Klon λ St 550-5 war es bei vier Hind III Stellen im Be-

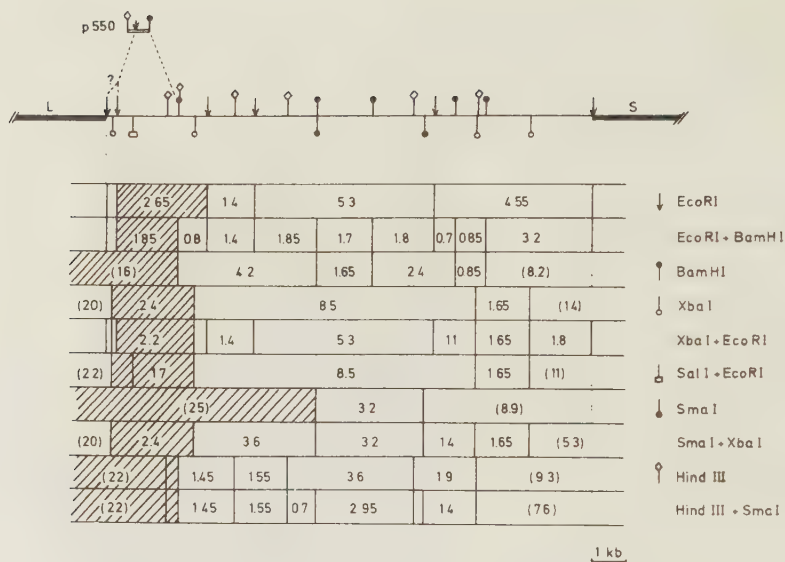


Abb. 23 : Restriktionskarte von λ St 550-3
 (Erklärungen s. Abb. 22)
 --? : Zuordnung p550 - Genom nicht eindeutig

reich des 5.3 kb Eco RI Fragments nicht möglich, diese genauer anzuordnen. Die Gesamtlänge der klonierten Inserts beträgt für den Klon λ St 550-3 nur 14.2 kb, für λ St 550-5 etwa 17.1 kb.

Der cDNA Klon p550 hybridisierte mit dem 2.65 kb und 0.7 kb Eco RI Fragment des Klons λ St 550-5 und mit dem 2.65 kb Fragment von λ St 550-3. Der hybridisierende Bereich innerhalb des 2.65 kb Eco RI Fragments konnte durch Spaltung mit Bam HI auf ein 1.65 kb langes Eco RI - Bam HI Fragment weiter eingegrenzt werden. Ob das 0.3 kb Eco RI Fragment ebenfalls p550-Sequenzen enthält, ließ sich anhand dieser Versuche nicht entscheiden, da Fragmente mit einer Länge von weniger als 400 bp nicht mehr effektiv genug an die Filter gebunden wurden und sich deshalb einer möglichen Hybrid-

sierung entzogen. Im Bereich des 0.7 kb Eco RI Fragments von λ St 550-5 lie sich der hybridisierende Bereich nicht genau lokalisieren.

Restriktionskarte des p550 Gens

Die Restriktionskarten der beiden genomischen Klone λ St 550-3 und λ St 550-5 zeigten im Bereich der 2.65 kb und 0.3 kb Eco RI Fragmente ein identisches Spaltungsmuster. Dieser Bereich reprsentiert damit den berlappungsbereich der beiden Klone. Die Karten der beiden Klone lieen sich zu einer Gesamtkarte der bisher isolierten genomischen Region zusammenfassen (Abb. 24), die einen Bereich von etwa 28 kb Lange umfat.

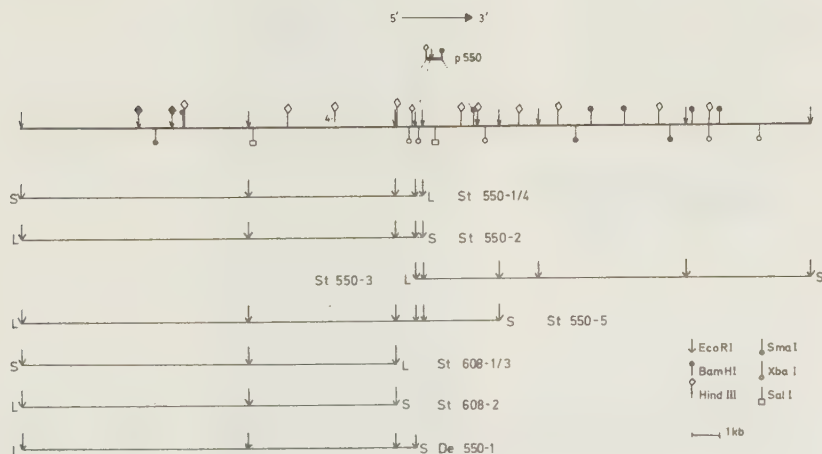


Abb. 24 : Zusammenfassende Restriktionskarte des p550-Gens und Lage der isolierten genomischen Klone
 L,S : Vektorarme von λ Ch 4A
 ... : Existenz des Fragments unsicher
 Klone : λ St 608-1,-2,-3 : s.C.4.3.
 λ De 550-1 : s.C.4.4.

Der cDNA Klon p550 enthält jeweils eine Eco RI-, Bam HI-, Hind III- und Sal I-Stelle. Zumindest die Bam HI- und Sal I-Spaltstelle ließen sich eindeutig den homologen Stellen im Genom zuordnen. Im Gegensatz dazu, finden sich im Falle der Eco RI- und Hind III-Stellen jeweils zwei alternative Möglichkeiten. Eine genauere Zuordnung war erst nach der Sequenzierung eines Teils dieser Region möglich. (s.C.5.).

Da die Orientierung der Sequenz des cDNA Klons p550 auf der mRNA bekannt ist (5'-Hind III-Bam HI-3'; s.C.1.3.), läßt sich auch die Richtung angeben, in der das 5'-Ende des p550-Gens und damit Promoter und regulatorische Sequenzen zu finden sein sollten. Die Richtung ist in Abb. 24 angegeben.

Ein Vergleich der Länge des hybridisierenden genomischen Bereichs (etwa 2.7 kb) und der Länge der cDNA (0.55 kb) führte zu dem Schluß, daß nur etwa 20% der Sequenz dieses Bereichs in der mRNA wiederzufinden sind. Damit muß das Gen in dieser Region sehr lange Intronbereiche enthalten. Falls dieses Verhältnis auf das gesamte Gen übertragbar ist, müßte für das gesamte Gen eine Länge von mindestens 20 kb (mRNA = 4.3 kb) angenommen werden. Es besteht somit die Möglichkeit, daß nicht das gesamte Gen in den isolierten Klonen enthalten ist.

Lage der isolierten genomischen Klone im p550-Gen

Neben der Restriktionskartierung der beiden Klone λ St 550-3 und -5 wurden auch alle anderen isolierten Klone kartiert. Es fanden sich dabei keine Unterschiede zu der bisher gefundenen Restriktionskarte. Aus diesen Daten ließ sich die relative Lage aller p550-spezifischen genomischen Klone angeben und deren Orientierung bezüglich der Vektorarme angeben. Eine Zusammenfassung ist in Abb. 24 gezeigt.

Auffallend bei der Verteilung der Klone ist die Häufung von Klonen, die das 8.1 und 5.3 kb Eco RI Fragment enthalten. Dies ist möglicherweise die Folge einer ungleichmäßigen Verteilung im Pool der Eco RI-Partialfragmente, die für die Konstruktion der Genbanken verwendet wurden.

3.5 Feinkartierung der Eco RI-Spaltstellen

Ein Problem bei der Kartierung der Eco RI-Spaltstellen war die richtige Einordnung des 0.3 kb Fragments. Alle Restriktionsspaltungen deuteten auf eine Lage zwischen dem 0.7 und dem 2.6 kb Eco RI Fragment hin. Eine exakte Bestimmung war wichtig, da in diesem Bereich p550-kodierende Sequenzen zu erwarten waren.

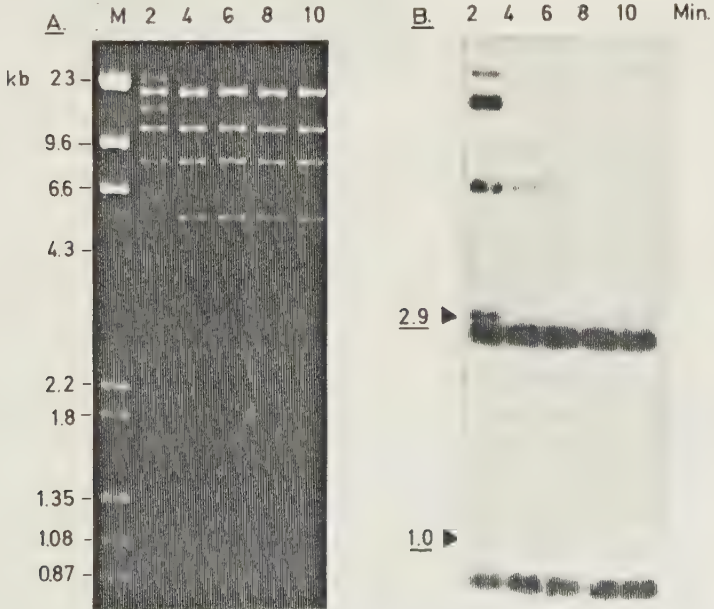


Abb. 25 : Lokalisierung des 0.3 kb/ Eco RI-Fragments
 A. Spaltung λ St 550-5 mit Eco RI (1U Eco RI/ μ g DNA) für verschiedene Zeiten (0.6% Agarosegel)
 B. Southerntransfer und Hybridisierung mit nick-translatiertem Insert von p550 (Autoradiographie)
 ►: Zusätzliche Partialfragmente

Eine Aussage über die Lage dieses Fragments war möglich durch die Partialspaltung des Klon λ St 550-5 mit Eco RI,

Transfer und Hybridisierung mit p550. Neben den 0.7 und 2.6 kb langen Fragmenten einer vollständigen Spaltung sollten noch zusätzliche hybridisierende Partialfragmente gefunden werden, die Kombinationen mit benachbarten Eco RI Fragmenten darstellen.

Das Ergebnis eines solchen Versuchs ist in Abb. 25 gezeigt. Zwei zusätzliche Fragmente mit Längen von 2.9 und 1.0 kb treten neben den erwarteten Banden auf. Diese Fragmente sind nur erklärbar, wenn das 0.3 kb Fragment zwischen dem 2.6 und 0.7 kb Eco RI Fragment angeordnet ist. Die postulierte Lage in der Restriktionskarte konnte dadurch bestätigt werden.

3.6 Vergleich der Restriktionskarten von Genom und klonierter Region

Eine genaue Analyse der isolierten genomischen Klone ist erst dann sinnvoll, wenn gezeigt werden kann, daß sie die Verhältnisse im Genom exakt widerspiegeln. Die Klone dürfen also nicht durch Ligation unabhängiger Fragmente zufällig entstanden sein oder durch Rekombinationsereignisse verändert worden sein. Allein die Tatsache, daß alle isolierten genomischen Klone identische Restriktionskarten zeigten, spricht gegen eine solche Möglichkeit.

Dennoch wurde diese Fragestellung durch Southernblotting von Spaltungen genomischer DNA und die Hybridisierung mit den radioaktiv markierten Klonen λ St 550-3 und -5 genauer überprüft. Alle Restriktionsfragmente der genomischen Klone sollten dabei wiedergefunden werden, neben zusätzlichen Randfragmenten, die über den bisher klonierten Bereich hinausreichen. Damit ließ sich auch prüfen, ob eventuelle Unterschiede im p550-Gen verschiedener Gewebe zu finden sind.

Zu diesem Zweck wurde DNA aus Sterna, Kalvarien (jeweils 17 Tage alte Embryonen) und von dedifferenzierten Chondrocyten (6 Wochen Kultur) mit Eco RI, Bam HI und Hind III ge-

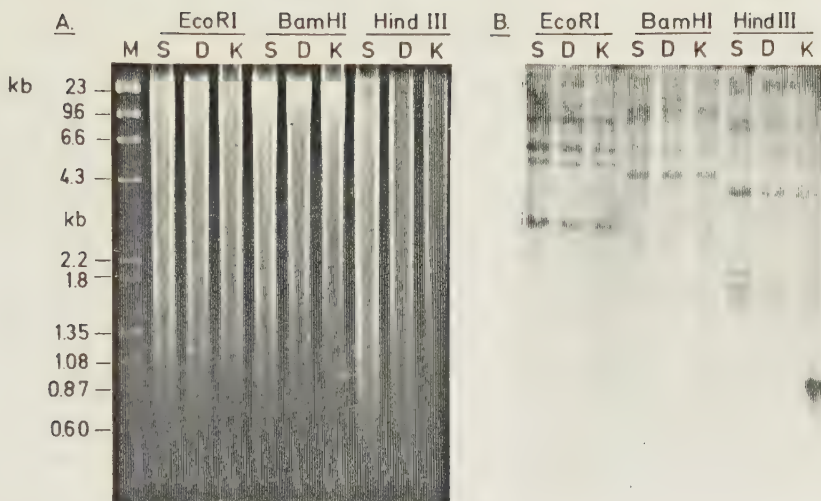


Abb. 26 : Vergleich der Restriktionskarten des p550-Gens im Genom mit den genomischen Klonen

A. Restriktionsspaltungen von 10 μ g genomischer DNA (0.8% Agarosegel)

B. Southerntransfer und Hybridisierung mit markierten λ St 550-3 und -5 (Autoradiographie)

S : Sterna-DNA

D : DNA aus dedifferenzierten Chondrocyten

K : Kalvarien-DNA

M : λ / Hind III + ϕ X/ Hae III

Tabelle 8 : Hybridisierung genomischer Fragmente mit den Klonen λ St 550-3 und -5 (Angaben in kb)

Fragmentlängen der Klone : s. Abb. 22 und 23

* : Randfragmente, die bisher nicht klonierte Anteile enthalten

(-3/-5) : Zuordnung zu λ St 550-3 oder -5

| Klone
Eco RI | Genom
Eco RI | Genom
Bam HI | Genom
Hind III |
|-----------------|-----------------|-----------------|-------------------|
| 8.1 | 8.1 | 11 | 15 |
| 5.3/5.3 | 5.3/5.3 | 8.3*(-5) | 8.5*(-3) |
| 4.5 | 4.5 | 5.5*(-3) | 7.5*(-5) |
| 2.5 | 2.6 | 4.2 | 3.7/3.6 |
| 1.3 | 1.3 | 2.4 | 1.9 |
| 0.7 | 0.7 | 1.65 | 1.6 |
| | | | 1.5 |
| | | | 1.45 |
| | | | 1.0 |

spalten, in einem Agarosegel getrennt und auf ein Filter transferiert. Bei der Auftrennung im Agarosegel ließen sich keine Banden erkennen, da Fragmente der verschiedensten Längen bei der Spaltung entstanden (Abb. 26A). Um eine Probe zu erhalten, die den gesamten klonierten Bereich repräsentiert, wurde die DNA von λ St 550-3 und -5 vollständig durch Nick Translation markiert. Eine gleichmäßige Markierung der Klone wurde dadurch erzielt, daß die Phagen-DNA vor der Markierung partiell mit Alu I und Hae III gespalten wurde. Die Filter wurden danach mit dieser Probe hybridisiert. Das Ergebnis einer solchen Hybridisierung zeigt Abb. 26B.

Die Restriktionsmuster der drei verwendeten Gewebe zeigten keine erkennbaren Unterschiede. Die Längen der hybridisierenden Banden sind in der Tabelle angegeben. Einige schwach erkennbare zusätzliche Banden sind wahrscheinlich auf eine unvollständige Spaltung zurückzuführen. Sehr große hybridisierende Banden (über 25 kb), die bei einigen Spaltungen zu finden sind, entsprechen noch ungespaltener DNA. Aus diesen Daten ließ sich kein Hinweis auf eine Variation in der Genomstruktur der unterschiedlichen Gewebe ableiten.

Das Fehlen deutlich hybridisierender, zusätzlicher Banden deutet auf das Fehlen von repetitiven Elementen innerhalb des klonierten Bereichs hin. Solche Sequenzen lassen sich in anderen Organismen in vielen genomischen Regionen nachweisen (88). Aufgrund dieses eindeutigen Hybridisierungsmusters ließ sich schließen, daß der Bereich der Klone λ St 550-3 und -5 nur Sequenzen enthält, die einmal im Genom auftreten.

Die Längen der Fragmente entsprachen den bei der Kartierung von λ St 550-3 und -5 gefundenen Werten (s. Abb. 22 und 23). Bei den durch ihre Länge nicht unterscheidbaren Fragmenten (5.3 kb Eco RI und 3.7 kb Hind. III) fanden sich nach der Hybridisierung intensivere Signale und entsprachen damit ebenfalls dem zu erwartenden Muster. Jeweils zwei zusätzliche Banden (*) wurden bei der Spaltung mit Bam HI und Hind III nachgewiesen. Sie repräsentieren terminale Restriktionsfragmente, die Bereiche enthalten, die in den bisher isolierten Klonen nicht enthalten waren. Durch zusätzliche

Hybridisierungen konnten diese Fragmente den einzelnen Klonen (λ St 550-3 oder -5) zugeordnet werden.

Diese Hybridisierungen zeigten, daß die Restriktionskarte der genomischen Klonen mit der des Genoms übereinstimmt und keine Unterschiede des p550-Gens in den verschiedenen (getesteten) Geweben existieren.

3.7 Methylierungsmuster des klonierten Genbereichs

Das Hauptinteresse bei der Charakterisierung eines Gens ist nicht nur die lineare Beschreibung der Genomregion und seiner Exon-Intron Struktur, sondern die Analyse der für die Regulation der Expression verantwortlichen Bereiche. Eine Steuerung der Genaktivität ist nicht nur möglich durch die Erkennung bestimmter Strukturen des Promoterbereichs (95), sondern auch durch Änderungen in Zustand und Struktur der DNA. Diskutiert werden dabei Veränderungen der Chromatinstruktur (96), der Helixtyptransitionen (97) und der spezifischen Methylierung der DNA (98).

Bei Eukaryonten findet sich die Methylierung der DNA fast ausschließlich in Form von 5-Methyl-Cytosin in der Sequenzfolge 5'-^mCG. Eine solche Methylierung scheint allgemein in einem umgekehrten Verhältnis zur Aktivität des Gens zu stehen, aber alleine noch keinen ausreichenden Faktor der Genregulation zu repräsentieren (99). Die Lage eines Teils der methylierten CG-Sequenzen läßt sich durch Spaltung genomischer DNA mit den beiden Isoschizomeren Hpa II und Msp I erkennen (100), von denen nur Msp I die methylierte Erkennungssequenz 5'-C^mCGG spalten kann.

Um die nicht methylierten CCGG-Sequenzen im Bereich des p550 Gens nachzuweisen, wurde deshalb die DNA von Sterna, Kalvarien und dedifferenzierten Chondrocyten mit Hpa II und Msp I gespalten, in einem Agarosegel aufgetrennt, auf Filter transferiert und mit der markierten DNA von λ St 550-3 und -5 hybridisiert.

Das Ergebnis ist in Abb. 27 gezeigt. Die Spaltung mit Msp

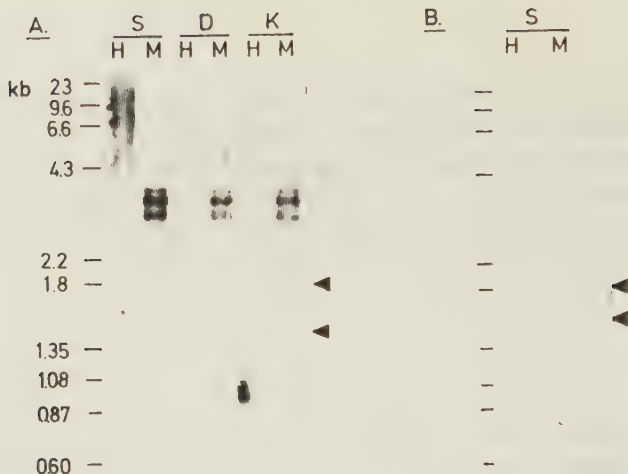


Abb. 27 : Methylierung des p550-Gens

A. Spaltungen von je 10 μ g genomischer DNA mit Hpa II (H) und Msp I (M), Southerntransfer und Hybridisierung mit markierten λ St 550-3 und -5 (Autoradiographie)

B. Hybridisierung mit markiertem Insert von pSt 608 (s.Abb.28)

▶: Nichtmethylierte Hpa II-Fragmente

S : Sterna-DNA

D : DNA aus dedifferenzierten Chondrocyten

K : Kalvarien

Es gibt das Muster aller CCGG-Sequenzen wieder, wobei sich wiederum kein erkennbarer Unterschied zwischen den Geweben erkennen läßt. Die meisten der Fragmente haben eine Länge von etwa 3 - 3.7 kb, und die Intensität der Banden weist auf eine große Anzahl gleich langer Fragmente hin. Bei einer absolut statistischen Verteilung der Basen sollten Spaltstellen für die Sequenz CCGG jedoch nach jeweils 256 bp ($1/4^4$) gefunden werden können. Diese Sequenzfolge liegt anscheinend unterrepräsentiert im Genom vor.

Bei der Spaltung mit Hpa II konnten nur zwei der Fragmente der Msp I Spaltung (1.7 und 1.9 kb) wiedergefunden werden. Das bedeutet, daß mit der Ausnahme von zumindestens

drei Positionen alle CCGG-Sequenzen in methylierter Form vorliegen. Nach längerer Exposition ließ sich zeigen, daß diese zwei Fragmente in allen Geweben zu finden sind, also auch in dieser Beziehung kein Unterschied besteht.

Eine Lokalisierung dieser Fragmente war durch die Hybridisierung mit Proben kleinerer Regionen der genomischen Klonen möglich. Dazu wurde neben verschiedenen anderen Fragmenten auch der Subklon pSt 608 (s.C.4.1.) verwendet, der dem 5'-terminalen Ende des klonierten Bereichs entspricht. Wurde mit dem markierten Insert von pSt 608 hybridisiert waren die 1.9 und 1.7 kb langen Fragmente sowohl bei der Spaltung mit Hpa II, als auch bei der Spaltung mit Msp I zu erkennen (s. Abb. 27B). Die nicht methylierten Hpa II Stellen ließen sich damit dem Bereich des Subklons pSt 608 zuordnen. Die genaue Lage dieser Hpa II-Stellen konnte durch die Restriktionskartierung und Sequenzierung dieses Subklons (s.C.6.) bestimmt werden. Eine mögliche Bedeutung dieser Positionen wird im Zusammenhang mit der Struktur des gesamten Gens diskutiert (s.G.4.).

4. Subklonierung und Isolierung weiterer genomischer Klone

4.1 Subklonierung verschiedener Genbereiche

Die in den λ -Vektoren klonierten Genfragmente sind zu lang, um eine einfache und detaillierte Untersuchung einzelner Regionen zuzulassen. Zu diesem Zweck ist es deshalb günstig, kleinere Fragmente in Plasmid- oder Phagenvektoren zu klonieren. Eines der Hauptziele war die Subklonierung terminaler Fragmente, mit deren Hilfe die Isolierung weiterer überlappender genomischer Klone möglich sein sollte. Der bekannte Genbereich könnte durch ein solches 'Walking up'-Experiment erweitert werden.

Die Genbank wurde konstruiert durch die partielle Spaltung der genomischen DNA mit Eco RI und die Ligation in λ Ch 4A (s.C.2.1.). Überlappende genomische Klone müssen damit zumindest die terminalen Eco RI-Fragmente des bisher klonierten Bereichs enthalten. Aus diesem Grund ist es ausreichend, beliebige Teilfragmente dieser Eco RI-Fragmente zur Suche nach überlappenden genomischen Klonen auszuwählen.

Die Lage der Fragmente ist in Abb. 28 gezeigt. Ein Eco RI/ Hind III-Fragment mit 5.8 kb Länge, das einen Teil des 5'-terminalen 8.1 kb Eco RI Fragments enthält, und ein Eco RI - Hind III-Fragment von 3.8 kb Länge im 3'-terminalen 4.55 kb Eco RI Fragment wurden zu diesem Zweck nach der Spaltung der entsprechenden λ -Klone (s.Tabelle 7) isoliert. Als Vektor wurde in beiden Fällen pBR 322 verwendet, der zuvor mit Eco RI und Hind III gespalten worden war. Vektor und isolierte Fragmente wurden ligiert und in HB 101 transformiert. Der Einbau eines Inserts zwischen die Eco RI- und Hind III-Stelle von pBR 322 führte zu einem Ausfall des Tetracyclinresistenzgens. Rekombinante Klone ließen sich

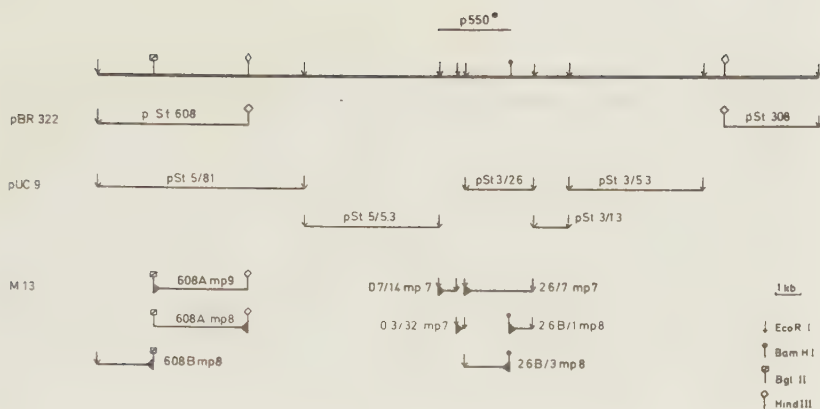


Abb. 28 : Restriktionskarte des p550-kodierenden Genbereichs und Lage der klonierten Fragmente einzelner Subklone
 pBR 322, pUC 9, M 13 : Verwendete Vektoren
 ► : Sequenzierungsrichtung der M13-Klone

deshalb an ihrer genetischen Konstitution Amp^res/Tet^sens erkennen. Der Stamm HB 101 wurde verwendet, da durch die $rec A^-$ -Mutation dieses Stamms eventuelle Rekombinationsereignisse innerhalb des klonierten Inserts verhindert werden.

Unter den Ampicillin-resistenten Transformanten fanden sich 60% (5.8 kb Fragment) bzw. 80% (3.8 kb Fragment) Tetracyclin-resistente Kolonien, die das gewünschte Fragment eingebaut enthielten. Einzelne Klone wurden ausgewählt und werden in der Folge als pSt 608 und pSt 308 bezeichnet. Beide Klone wurden später zum erneuten 'Screening' der Genbank verwendet (s.C.4.3.).

Die meisten Eco RI Fragmente der beiden Klone λ St 550-3 und -5 wurden ebenfalls in einem Plasmidvektor subkloniert. Um die Selektion rekombinanter Klone zu erleichtern, wurde dazu der Vektor pUC 9 (101) verwendet, da in pBR 322 der Einbau von Eco RI Fragmenten in pBR 322 zu keiner Veränderung der Selektionsmarker führen wurde. Der Vektor pUC 9 enthält dagegen neben Teilen von pBR 322 (Amp^res) einen Teil des lac-Gens von E.coli. Ähnlich den M13 Vektoren, ist auch

hier eine Polylinkerregion enthalten (s. Abb. 7), die eine Reihe von Klonierungsstellen (darunter Eco RI) aufweist. Der Einbau eines Eco RI-Fragmente läßt sich wiederum am Ausfall der β -Galaktosidase Aktivität, in einem geeigneten Empfängerstamm (JM 83), erkennen. Rekombinante Klone sind deshalb an der Resistenz gegenüber Ampicillin zu erkennen, und die Kolonien zeigen auf X-Gal-haltigen Platten keine Blaufärbung. In allen Fällen ließen sich nach der Ligation des Vektors mit den isolierten Fragmenten und der Transformation Subklone erhalten, die das jeweils gewünschte Fragment enthielten.

Tabelle 7 : Subklone der p550 kodierenden Genregion

| Klon | Herkunft | Fragment-
länge | Klonierungs-
stelle | Vektor |
|------------|--------------------|--------------------|------------------------|---------|
| pSt 608 | λ St 550-5 | 5.8 kb | Eco/Hind | pBR 322 |
| pSt 308 | λ St 550-3 | 3.8 | Eco/Hind | " |
| pSt 5/8.1 | λ St 550-5 | 8.1 | Eco | pUC 9 |
| pSt 5/5.3 | λ St 550-5 | 5.3 | Eco | " |
| pSt 3/2.6 | λ St 550-3 | 2.65 | Eco | " |
| pSt 3/1.3 | λ St 550-3 | 1.3 | Eco | " |
| pSt 3/5.3 | λ St 550-3 | 5.3 | Eco | " |
| 608A mp9 | pSt 608 | 3.7 | Hind/Bam-Bgl | M13 mp9 |
| 608A mp8 | pSt 608 | 3.7 | Hind/Bam-Bgl | " mp8 |
| 608B mp8 | pSt 608 | 2.1 | Eco/Bam-Bgl | " mp8 |
| 0.7/14 mp8 | λ St 550-5 | 0.7 | Eco | " mp7 |
| 0.3/32 mp7 | λ St 550-5 | 0.3 | Eco | " mp7 |
| 2.6/7 mp7 | λ St 550-3 | 2.65 | Eco | " mp7 |
| 2.6B/1 mp8 | λ St 550-3 | 0.8 | Eco/Bam | " mp8 |
| 2.6B/3 mp8 | λ St 550-3 | 1.65 | Eco/Bam | " mp8 |

Neben diesen Subklonen in Plasmidvektoren wurde eine Reihe von Fragmenten ebenfalls in M13-Vektoren eingebaut.

Die Lage dieser Fragmente und deren Orientierung nach dem Einbau in M13 ist in Abb. 28 zusammengefaßt. Zur Klonierung der Eco RI-Fragmente wurde M13 mp7 verwendet. Um eine definierte Einbaurichtung der Fragmente mit unterschiedlichen Enden zu erreichen, wurden diese in die beiden Vektoren M13 mp8 und mp9 kloniert. Zur Subklonierung von Fragmenten des Subklons pSt 608 wurde die dort nur einmal auftretende Bgl II-Stelle (T/GATCA) verwendet. Die bei der Spaltung auftretenden Einzelstrangenden (GATC-) erlauben eine Klonierung in die Bam HI-Stelle (G/GATCC), wobei eine hybride Spaltstelle (GGATCA) entsteht, die von keinem der beiden Enzyme nach der Ligation gespalten werden kann.

Herkunft der klonierten Fragmente, die verwendeten Vektoren und Klonierungsstellen sind in der Tabelle 7 für alle Subklone zusammengefaßt.

4.2 Eingrenzung des kodierenden Bereichs

Aus der Kodierungskapazität im Bereich der p550-Sequenz von etwa 20% (s.C.3.4.) ergibt sich bei einer gleichmäßigen Verteilung der Exons über das gesamte Gen eine Gesamtlänge von mindestens 20 kb (mRNA etwa 4.3 kb). Falls p550 eine zentrale Sequenz innerhalb der mRNA repräsentiert, ist es möglich, daß bisher nicht das komplette Gen isoliert wurde. Der Genbereich ließ sich eingrenzen durch die Hybridisierung von RNA mit verschiedenen Subklonen. Die beiden terminalen Klone pSt 308 und pSt 608 eigneten sich besonders gut, da ihre Größe es wahrscheinlich machte, daß neben Intron- auch Exon-Sequenzen enthalten sind, die eine Hybridisierung mit mRNA möglich machen.

Je 10 µg RNA, isoliert aus *Sterna* und Kalvarien (17 Tage alte Embryonen), wurden mit Glyoxal denaturiert, in einem 1% Agarosegel aufgetrennt und auf ein Filter transferiert. Zur Hybridisierung wurden vier verschiedene radioaktiv markierte Proben verwendet, deren Lage in der Abb. 29 angegeben ist. Durch Nicktranslation wurde das Insert 3'-terminalen Sub-

klons pSt 308 und das 5'-terminale 2.2 kb lange Eco RI/Bgl II-Fragment (608 B) des Subklons pSt 608 markiert. Der M13 Klon 608A mp9 wurde markiert, da in einem Vorversuch gezeigt wurde, daß nur dieser Klon und nicht der komplementäre Klon 608A mp8 mit RNA hybridisiert.

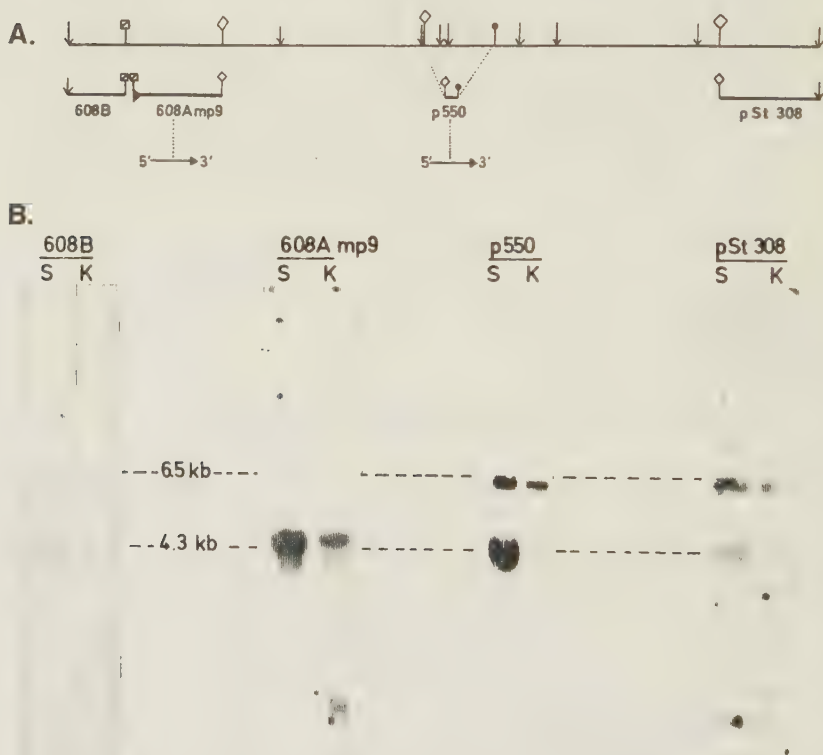


Abb. 29 : Northernhybridisierungen mit verschiedenen Proben des p550-Genbereichs
 A. Restriktionskarte des p550-Gens und Lage der verwendeten Hybridisierungsproben
 B. Northerntransfer von 10 µg RNA und Hybridisierung mit den angegebenen Proben (Autoradiographien)
 RNA : S = Sterna/ 17 Tage; K = Kalvarien/ 17 Tage
 608B*, p550*, pSt 308* : Nicktranslation der isolierten Fragmente
 608A mp9* : Markierung des M13 Klons

In Abb. 29 ist das Ergebnis dieser Hybridisierungen zusammengefaßt. Der 3'-terminale Klon pSt 308 wies ein identisches Hybridisierungsmuster zu dem des cDNA Klons p550 auf. Das 3'-terminale Ende des p550 Gens ist damit entweder bereits in diesem Bereich zu suchen, kann jedoch auch in einem bisher nicht erfaßten Bereich außerhalb der isolierten Klone liegen. Die Hybridisierung mit den beiden 5'-terminalen Fragmenten (608A und 608B) ergab dagegen ein überraschendes Ergebnis.

Das 5'-terminale 2.1 kb Eco RI/ Bgl II-Fragment (608B) zeigte wiederum eine Hybridisierung mit den 4.3 und 6.5 kb langen mRNA Species. Im Gegensatz dazu ließ sich aber mit 608A mp9, das den direkt anschließenden genomischen Bereich erfaßt, eine Hybridisierung ausschließlich mit der 4.3 kb mRNA nachweisen. Aus der Einbaurichtung des Inserts in diesem Klon ließ sich die Orientierung der mRNA in diesem Bereich schließen: 5'-Bgl II-Hind III-3'; sie entspricht damit der Orientierung des cDNA Klons p550. Durch zusätzliche Hybridisierungen konnte gezeigt werden, daß das Verteilungsmuster von 608A mp9 mit RNA verschiedener Gewebe dem von p550 sehr ähnlich ist.

Anhand dieser Ergebnisse sind verschiedene Rückschlüsse auf das p550 Gen möglich. Die Ähnlichkeit der Hybridisierungsmuster macht es wahrscheinlich, daß sowohl pSt 308 als auch pSt 608 Teile des p550 Gens enthalten. Dafür spricht neben den vergleichbaren Längen der mRNAs auch die parallele Orientierung von p550 und pSt 608. Der kodierende Bereich umfaßt zumindest den gesamten klonierten Bereich, also etwa 25 kb. Eine genaue Festlegung der Grenzen des Gens ist damit noch nicht möglich. Die theoretische Möglichkeit, daß zwei benachbarte, unabhängige Gene zur Synthese von identisch langen mRNAs führen, ist wenig wahrscheinlich.

Die Frage, in welchem Verhältnis die beiden mRNA-Species stehen, wurde bereits früher aufgeworfen (s.C.1.3.) und wird durch diese Befunde noch weiter kompliziert. Zu einer genauen Beschreibung der Expression, des Splicings und der kodierten Proteine (s.G.5.) reichen die bisherigen Befunde

noch nicht aus.

4.3 Versuche zur Ausweitung des bisher isolierten Genbereichs

Da die Experimente nahelegten, daß das Gen bisher noch nicht vollständig erfaßt worden war, sollte mit Hilfe terminaler Subklone in einem 'Walking Up'-Experiment der klonierte Genbereich ausgeweitet werden. Die beiden Subklone pSt 308 und pSt 608 waren primär zu dem Zweck konstruiert worden, um aus der Genbank überlappende Klone isolieren zu können. Ähnlich dem Screening der Genbank mit p550 (s.C.3.2.) wurden 200.000 pfu der primären 'Sterna'- Genbank plattiert, auf Filter transferiert und für die Hybridisierung vorbereitet. Zur Hybridisierung wurden die Inserts von pSt 308 und pSt 608 durch Nicktranslation markiert und gleichzeitig zur Hybridisierung eingesetzt. Dabei konnten drei hybridisierende Klone nachgewiesen werden.

Um entscheiden zu können, welche der beiden Proben hybridisierte, wurde bei der ersten Reinigung der positiven Plaques die doppelte Anzahl paralleler Filter hergestellt und jeweils zwei der vier Filter mit pSt 308 bzw. pSt 608 hybridisiert. Bei allen drei isolierten Klonen fand sich dabei nur eine Hybridisierung mit pSt 608, nicht jedoch mit pSt 308. Es konnte damit kein weiterer, für das 3'Ende des Gens spezifischer Klon isoliert werden. Die drei Klone wurden bis zur Homogenität gereinigt und mit λ St 608-1/-2 und -3 bezeichnet.

Restriktionskartierung der Klone λ St 608-1/-2/-3

Die Restriktionskartierung der Klone λ St 550-1, -2 und -3 sollte zeigen, ob in diesen Klonen bisher nicht klonierte Bereiche enthalten sind. Die isolierte DNA wurde deshalb mit Eco RI und Bam HI gespalten und die Längen der entstandenen Fragmente sind in der Tabelle 9 angegeben.

Bei der Spaltung mit Eco RI fanden sich in allen Fällen nur zwei Fragmente von 5.3- und 8.1 kb Länge, die genau den Fragmenten des Klon λ St 550-1 entsprachen. Offensichtlich enthielten alle drei Klone identische Inserts von etwa 13.4 kb Länge, die keine zusätzlichen Bereiche außerhalb der bisher klonierten Region beinhalten.

Tabelle 9 : Restriktionsfragmente der Klone λ St 608-1/-2/-3

* : Fragmente mit Teilen von Vektor und Insert

() : Fragmente der Vektorarme

| | | | | | | |
|----------------------|-------------|-------|-------------|-------|-------------|-------|
| Klone : | λSt 608-1 | | λSt 608-2 | | λSt 608-3 | |
| Spaltung : | Eco | Bam | Eco | Bam | Eco | Bam |
| Fragmente : | (19) | 22* | (19) | 20* | (19) | 22* |
| (kb) | (12) | 11* | (12) | 12* | (12) | 11* |
| | 8.1 | (5.5) | 8.1 | (5.5) | 8.1 | (5.5) |
| | 5.3 | (3.9) | 5.3 | (3.9) | 5.3 | (3.9) |
| | | (1.5) | | (1.5) | | (1.5) |
| Anordnung : | S-8.1-5.3-L | | L-8.1-5.3-S | | S-8.1-5.3-L | |
| der Eco RI Fragmente | | | | | | |

Die Anordnung der Fragmente ließ sich aus der Spaltung mit Bam HI erschließen, da sich nur eine Bam HI-Stelle innerhalb des 8.1 kb Eco RI-Fragments findet (s. Abb. 24). Aus der Länge der Fragmente, die sowohl Teile des Vektors als auch des Inserts enthielten, ließen sich die relative Orientierung des Inserts in Bezug zu den Armen des Vektors ableiten. In den beiden identischen Klone λ St 608-1 und λ St 608-3 ist im Vergleich zu λ St 608-2 das klonierte Insert in der entgegengesetzten Richtung integriert. Die Kartierung zusätzlicher Restriktionsstellen bestätigte die Anordnung der Fragmente und die Zuordnung zum bereits bekannten Genbereich (s. Abb. 24).

Der Bereich in der Umgebung des Subklons pSt 608 wird da-

mit durch insgesamt 7 genomische Klone repräsentiert, während die Region des Subklons pSt 308 nur in einem genomischen Klon wiedergefunden wurde (Abb. 24). Diese Verteilung spiegelt wahrscheinlich die ungleiche Verteilung einzelner Bereiche in den klonierten Fragmenten der Genbank wider.

4.4 Isolierung des p550-Gens aus der DNA dedifferenzierter Chondrocyten

Auch in der 'Dediff.'-Genbank (s.C.2.), hergestellt aus der DNA dedifferenzierter Chondrocyten, wurde nach Klonen gesucht, die Teile des p550-Gens enthalten. Beim Screening von 200.000 pfu der primären Genbank durch Hybridisierung mit dem markierten Insert von p550 konnte nur ein positiver Klon identifiziert werden: λ De 550-1. Die Restriktionskartierung dieses Klons ergab, daß er Eco RI-Fragmente von 8.1, 5.3 und 0.7 kb Länge enthielt. Durch die Spaltung mit weiteren Enzymen ließ sich zeigen, daß dieser Klon den bereits isolierten Klonen λ St 550-1/-2 und -4 entsprach (s. Abb. 24). Da dieser Klon keine signifikanten Unterschiede zu der bisher gefundenen Restriktionskarte aufwies, bestätigte dies das Ergebnis der genomischen Hybridisierungen (s.C.3.6.), wonach keine Unterschiede in der Struktur des p550-Gens in verschiedenen Geweben - zumindest zwischen Sternum, Kalvarien und dedifferenzierten Chondrocyten - zu finden sind.

5. Feinanalyse des p550 kodierenden Genbereichs

5.1 Sequenzierung des p550 kodierenden Genbereichs

Bei der Kartierung der genomischen Klone konnte die p550-kodierende Region auf einen Bereich von etwa 2.6 kb Länge eingeschränkt werden (s.C.3.4.). In Abb. 30 ist die Restriktionskarte dieser Region gezeigt.

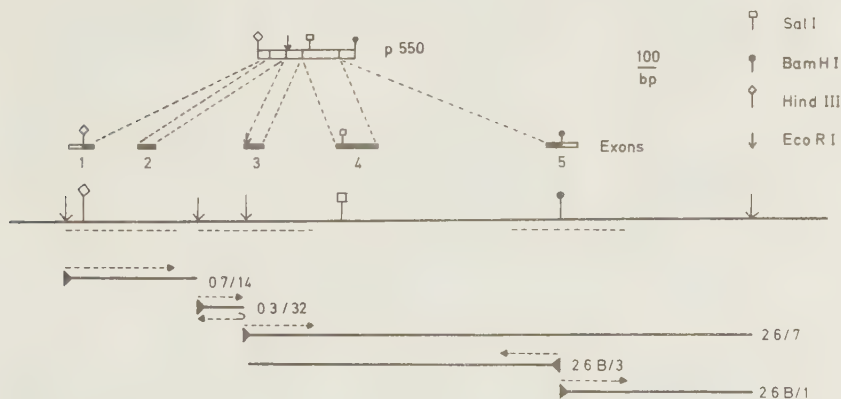


Abb. 30 : Sequenzierung des p550-kodierenden genomischen Bereichs und Lage der verwendeten M13-Klone

- : Exonbereiche
- ▭ : Hypothetische Exonbereiche
- > : Sequenzierte Bereiche
- <--- : 'Reverse' Sequenzierung

Zumindest die Sal I- und Bam HI-Spaltstellen des cDNA-Klons konnten den entsprechenden Stellen des Genoms eindeutig zugeordnet werden. Die Eco RI- und Hind III-Stellen konnten aber erst nach der Sequenzierung dieser Region

Strang nach der Didesoxymethode sequenziert werden. Dazu wurde ein Primer (17 b; 'Reverse Sequencing Primer'; s.G.2.10.) verwendet, der vor dem 3'-Ende des Inserts im (-)-Strang bindet. Die auf diese Weise erhaltene Sequenz war vollkommen komplementär zu der vorher bestimmten Sequenz der Gegenrichtung, sodaß eine eindeutige Sequenz für diesen Bereich festgelegt werden konnte.

Anhand der Restriktionskarte ließen sich diese Sequenzen zu einer Gesamtsequenz zusammenfügen, die in Abb. 32 gezeigt ist. Die Sequenz ist bereits in der Orientierung des kodierenden Strangs, also 5'-Hind III-Bam HI-3' (s.C.1.3.), angegeben, um einen direkten Vergleich mit der cDNA-Sequenz zu erleichtern. Die Längen der nicht sequenzierten Bereiche wurden aus den Daten der Restriktionskartierung abgeschätzt; um die Orientierung innerhalb der Gesamtsequenz zu erleichtern, wurde die Nummerierung der Basen darauf abgestimmt.

Abb. 32 : Sequenz der p550-kodierenden genomischen Region

```

..... : Nicht sequenzierte Bereiche (Längenangaben entsprechen Abb. 23)
NNNN   : Restriktionsstellen : Eco RI, Hind III
                               Bam HI
????   : Mögliche Exon-Intron Übergänge
****   : Unterschiede zur p550-Sequenz (s. Abb. 11)
NNNN  : Sequenz identisch p550
!       : Beginn/ Ende der cDNA Sequenz
NNNN   : Unterschiedliche Sequenz in p550

```

1 GAATTC²²²²CCCTG CAGGAGATCT TNGAGAAAAC TGGAGTCTCT GTCGAGATCC
51 CACCCACTGA CAGTAGCTCG GAGACGGTGA TACTGCGAGG CGAGCCCGAA
101 AAGCTTGGGC AAGC^{*}TTGAC TGAAGTCTAT GCAAAGGTAT CGATGAGATG
151 GGCTGGGTCA GATGGGTAA AAAGTCACGT TAGGATATAT GTGGGAGCAA
201 CATGGACAAG TGGAGGAGGG GTGGTCTTTG TTGGGCTGTT AGAAATGCTG
251 CTTAGATTAA CTCAATACTG AGGATACAAC ATAGAAAATTG GAACTGAATT
301 TTGTGAAGA CTTCTTCTGG TCCATCCATC AAAAAACAGCT TTTTAAATTT

351 TCTGTCTTTC AGGCCAACAG TTTTACCGTC TCCTCGGTCT CTGCCCCCTC
 401 TGGGCTTCAT CGTTTCATTA TTGG^{***}AA^{***}AA AGGTCAAAAC CTGGCCAAAA
 451 TA^{***}ACTCAGCA AATGCCAAAG GTGAGCT...
 (etwa 220 bp)
 701 GAATTCCCCG GTTGCGTCGA CCTGCACGAT NGTGAGATAA AGGTTAGAGA
 751 GGTGAGCAAY TYCTCTGAAG CCCTTGCAAA CAATAAGAAC AGCATGAGAG
 801 CCATAGGCTC TAGATTCTGT GACATGGCCA AAATGGGGCA TTACTGTCTT
 851 CTGAAAGCTC ATGAATAATG CGTTTGCTT GTGGGAAGGA AAATATCTTG
 901 TTAGTGGCAG AGATGATGTA GGCCTGACTA AGGGCATTCA TCTGTCATGC
 951 ATCTTGAAA ATCCTTAATG AATTC.... (20-30 bp?) ATTACATC
 1001 GAATTCAC^{***}TCG AAGGAGAGGA TAAATCACT TTGGAAGGAC CTACAGAAGA
 1051 TGTGAATGTG GCTCAGGAAC AGATTGAAGT CATGGTCAAG GATCTGGTAA
 1101 GTTGGTGTTA GTAGTTTCAG TTCATAAGAA GGGGCGATAG AACAGAGATG
 1151 GTAAGTTACC TATCACATTT ATGATGTTTT TNCATGCTTA ACAGAACTTA
 1201 CAGGAGGCAG TTNNTACGTC AGATGGTTAC AAAANTTNTC CACAGCANTC
 1251 TTA.....
 (etwa 200 bp)
 Exon 4 (219 bp/ p550 Pos. 250 - 468)
 (etwa 750 bp)
 2401 GGTTTTGCGC TGACAATTAA TTTGTTAGGC ATCAAATGGG TGAAATAAAA
 2451 TATATGGCTC GACTGCTGAA TACTCAATTG TGTATTACCA TCAGG^{***}AAAAAT
 2501 GAACGCACCA AGGACCTAAT CATIGAGCAG AAATTCACC GGACCATCAT
 2551 TGGGCAGAAG GGTGA^{***}GCGGA TC^{***}TA^{***}ACTCA CATTAATTGC G^{***}TTGCGCTCA
 2601 CTGCCCGCTT TCCA^{???}GT^{???}GGG AAACCT^{???}GT^{???}GC^{???}CT^{???}GC^{???} ATTAATGAA
 2651 TCGGCCAACG CGCGGGGAGA GCGGGTTTGC TATTGGGCG CCAGGGTGG
 2701 TTTTCTTTT CACCAGTGAG ACGGGCAACAG CTGATTGCC CTTACCGC
 2751 CTCCGCCCTG AGAGAGTTGC AGCAAGCGGT CCANGGTGGT TTGCCCCAGN
 2801 AGGNGAAAAAT CCTGTTTGAT GGTGGTT

5.2 Exon-Intron Struktur des p550 kodierenden Bereichs

Vergleich der Sequenzen von cDNA und Genom

Die Sequenz des cDNA Klons p550 findet sich in der genomischen Sequenz wieder und ist auf insgesamt 5 Exons verteilt. Die genaue Zuordnung der Sequenzen von p550 und der genomischen Sequenzen ist in der Abb. 32 angegeben und in der Tabelle 10 im Detail zusammengefaßt.

Tabelle 10 : Vergleich p550 - genomische Sequenz
(Positionen : p550 s.Abb. 11/ Genom s.Abb. 32)

| Exon | Exonlänge | p550 | Genom |
|------|-----------|---------|-----------|
| 1 | 36+ bp | 1-36 | 101-136 |
| 2 | 108 " | 37-144 | 363-470 |
| 3 | 105 " | 145-153 | ? |
| | | 154-249 | 1001-1096 |
| 4 | 219 " | 250-468 | n.b. |
| 5 | 79+ " | 469-547 | 2495-2573 |

Die Übereinstimmung der Sequenz von p550 mit der genomischen Sequenz beginnt an der Hind III-Stelle (Pos. 101), also innerhalb des Exons 1 (s. auch Abb. 31). Das 2. Exon mit einer Länge von 108 bp beginnt bei Pos. 363 im Genom. Der dem Exon 4 entsprechende Bereich wurde nicht sequenziert. Da die Sal I Spaltstelle des cDNA Klons jedoch einer entsprechenden Stelle im Genom zugeordnet werden kann, sollte dieses Exon etwa zwischen Position 1450 und 1670 zu finden sein. Einen Hinweis darauf, daß diese Sequenz nur innerhalb eines Exons vorliegt, gaben R-Loop Experimente (102). Dazu wurde mRNA mit der DNA von 550 mp8 hybridisiert und die entstehenden Hybride im Elektronenmikroskop sichtbar gemacht.

Ein längeres Exon konnte in einem Bereich identifiziert werden, der dem des Exons 4 entsprach. Das Exon 5 beginnt mit der Pos. 2495 und reicht noch über die Bam HI Stelle (Pos. 2573) hinaus, die dem Ende der cDNA Sequenz entspricht.

Probleme in der Zuordnung der cDNA und der genomischen Sequenz gab es nur in der Umgebung der Eco RI Spaltstelle von p550. Aus der Restriktionskarte konnte angenommen werden, daß diese Stelle einer der beiden Eco RI Spaltstellen des 0,3 kb Eco RI Fragments entspricht (Pos. 701 und 1001). Die Sequenz dieses Fragments (701 - 975) zeigte aber keinerlei Homologie mit p550. Das Exon 3 sollte jedoch mit der Sequenz ATTCACATC- (Pos.145-153) beginnen. Diese Folge von 9 bp wurde im Genom nicht wiedergefunden. Erst die darauf folgende Sequenz von p550, beginnend mit der Eco RI Stelle (Pos. 154), fand sich im angrenzenden 2,6 kb langen Eco RI-Fragment (Pos. 1001 - 1096). Da die Lage des 0,3 kb Eco RI Fragments an dieser Stelle der Restriktionskarte gesichert ist (s.C.3.5.), muß hier im Genom noch ein zusätzliches, sehr kurzes Eco RI Fragment angenommen werden, das bisher bei der Kartierung übersehen wurde. Ein Eco RI Fragment von etwa 20-30 bp Länge wäre bei den verwendeten Methoden der Kartierung nicht nachzuweisen gewesen. Andere Erklärungen - wie Fehler in der cDNA Sequenz oder ein zusätzliches Exon von 9 bp Länge - sind weniger wahrscheinlich. Durch die Subklonierung anderer, überlappender Fragmente kann die Sequenz dieses kleinen Fragments bestimmt werden.

Unterschiede zwischen der cDNA und der genomischen Sequenz finden sich nur an zwei Stellen: an Pos. 115 fand sich im Genom ein G anstelle eines A in p550; im Bereich 425-431 sind alle A's der genomischen Sequenz gegen T's in der p550 Sequenz ausgetauscht. Sonst traten keine Unterschiede zwischen der Sequenz von p550 und dem kodierenden Bereich des Genoms auf. Zwei Erklärungen bieten sich für diese Unstimmigkeiten an: entweder sind bei der Synthese der cDNA durch die Reverse Transkriptase Fehler eingeführt worden, oder die Unterschiede sind eine Folge der Klonierung in λ Ch 4A und M13. Auf welcher Ebene diese Unterschiede entstanden sind,

läßt sich anhand der Sequenzdaten allein nicht entscheiden.

Vergleich der Exon-Intron Übergänge

Durch den Vergleich der Sequenzen vieler Intron-Exon und Exon-Intron Übergänge sind Prototypsequenzen abgeleitet worden, die in ähnlicher Form bei vielen Genen gefunden wurden und wahrscheinlich eine wichtige Rolle beim Vorgang des Spleißens spielen (103). Diese Sequenzen sind in der Tabelle angegeben.

Tabelle 11 : Exon-Intron Struktur im Bereich der p550-kodierenden Sequenz des Genoms

() : Hypothetische Übergänge

* : Mehrere Alternativen möglich

| | | | | |
|---------------------|-----------------|-------------------|--------------------------|-----------------------------|
| Py _{6/9} | (GCAG - GAG)... | <u>Exon 1</u> ... | AAG - GTATCG
Lys | <u>Intron A</u>
(226 bp) |
| | | (123 bp) | | |
| Py _{10/11} | TCAG - GCC... | <u>Exon 2</u> ... | AAG - GTGAGC
Ala Lys | <u>Intron B</u>
(450 bp) |
| | | (108 bp) | | |
| | - ATT... | <u>Exon 3</u> ... | CTG - GTAAAGT
Ile Leu | <u>Intron C</u>
(350 bp) |
| | | (105 bp) | | |
| | - ATC... | <u>Exon 4</u> ... | ATG -
Ile Met | <u>Intron D</u>
(850 bp) |
| | | (219 bp) | | |
| Py _{6/11} | TCAG - GAA... | <u>Exon 5</u> ... | (CCT - GTCGTG)
Glu | |
| | | (132 bp) | | |

Prototypsequenz :

Py₁₁ N^CTAG - G C^GAG - GT^GAGT

Mit diesen Sequenzen konnten nun alle Übergänge im Bereich der p550 kodierenden Sequenz verglichen werden. Diese Prototypsequenzen wurden auch dazu verwendet, den Beginn des Exons 1 und das Ende des Exons 5 festzulegen. Neben diesem Kriterium der Sequenzhomologie sollte der Leseraster der p550-Sequenz auch in diese Sequenzbereichen fortgesetzt werden können, ohne Stopcodons aufzuweisen. Als hypothetischer

Beginn des Exons 1 wurde die Position 14 angenommen, das Ende des Exons 5 wurde mit Position 2627 festgelegt. Alternative Übergänge sind auch denkbar an den Positionen 2590, 2614 oder 2635.

Die Längen aller Exons, Introns und die Sequenzen der Übergänge sind in der Tabelle 11 zusammengefaßt. Ein Vergleich der Sequenzen läßt erkennen, daß die Prototypsequenzen in den meisten Fällen recht gut konserviert sind. Alle Übergänge folgen der AG-GT Regel (104), wonach ein Intron meist mit AG- endet, und das nächste Intron mit -GT beginnt.

Ein interessanter Aspekt ergibt sich, wenn man den Leseraster mit der Lage der Exon-Intron Übergänge in Beziehung bringt. An allen sequenzierten Übergängen liegt die Spaltstelle nie innerhalb eines Tripletts, sondern immer zwischen den Codons. Bei den meisten der bisher beschriebenen Genen findet sich kein solcher Zusammenhang zwischen Leseraster und Exon-Intron Übergängen.

6. Charakterisierung des 5'-terminalen Bereichs des p550 Gens

6.1 Restriktionskartierung des Klons pSt 608

Der Subklon pSt 608 enthält den am weitesten 5'-terminal gelegenen Bereich des p550-Gens, der in den isolierten genomischen Klonen enthalten ist. Aus zwei Gründen wurde dieser Bereich näher untersucht: Erstens war denkbar, daß dieser Klon bereits das 5'-Ende der mRNA enthielt und damit möglicherweise auch bereits Promotersequenzen. Das Vorhandensein nicht-methylierter Hpa II-Stellen in diesem Bereich des Genoms (s.C.3.7.), gab bereits einen Hinweis auf die Nähe eines Promoters (131). Ein zweiter interessanter Befund war das eigenartige Hybridisierungsmuster mit mRNA (s.C.4.2.), da Sequenzen des 3.7 kb Bgl II/ Hind III-Fragments (608A) im Gegensatz zu allen anderen getesteten genomischen Fragmenten nur in der 4.3 kb langen mRNA nachzuweisen waren, nicht jedoch in der 6.5 kb mRNA (s.Abb. 29).

Die Restriktionskarte des Klons pSt 608 wurde durch die Spaltung des gesamten Plasmids oder des isolierten Inserts mit verschiedenen Restriktionsenzymen bestimmt. Die zusammengefaßte Karte ist in Abb. 33 gezeigt. Die Lage der Spaltstellen einiger häufig spaltender Enzyme, wie Hpa II, Hae III, Hha I und Pst I, konnte nicht durch einfache Spaltungen des Klons bestimmt werden. Besonders die Lage der in großer Zahl vorhandenen Hpa II-Stellen war wichtig, da zumindest drei entsprechende Stellen im Genom in nicht methylierter Form vorliegen. Die Kartierung solcher Stellen war möglich durch die partielle Spaltung eines endmarkierten Fragments. Das Eco RI-Hind III Insert des Klons pSt 608 wurde dazu an den 5'-Enden mit $\gamma^{32}\text{P}$ -ATP und Polynukleotidki-

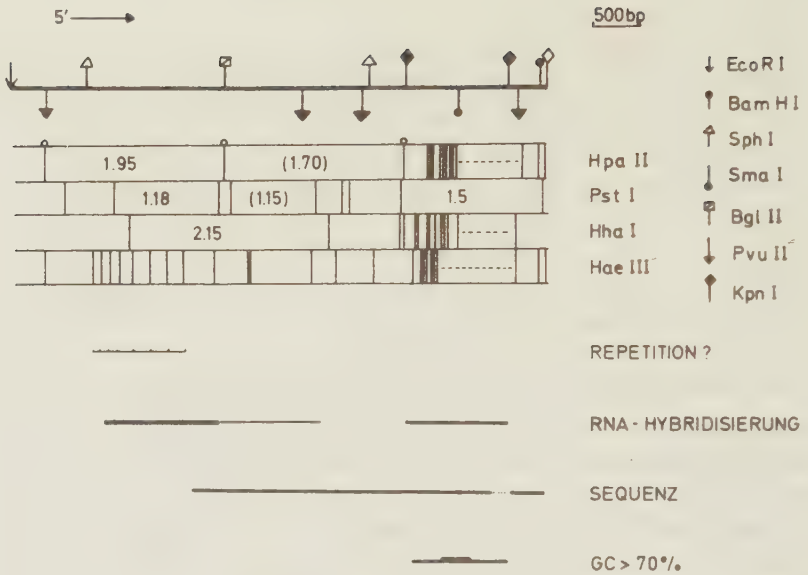


Abb. 33 : Restriktionskarte des klonierten Inserts des Subklons pSt 608 und Lage charakteristischer Regionen

---- : große Anzahl von Spaltstellen

(NN) : Fragmente, die mit markierter RNA hybridisieren (s.C.6.2.)

nase markiert und danach mit Bam HI gespalten. Dabei entstehen zwei, nur an einem Ende markierte Fragmente von 5.8 kb und 50 bp Länge. Das kurze Bam HI/Hind III-Fragment störte im weiteren Verlauf nicht das partielle Spaltungsmuster des großen Eco RI/Bam HI-Fragments. Dieses endmarkierte Fragment wurde unter verschiedenen Bedingungen (Enzymmenge, Inkubationszeit) mit einigen Restriktionsenzymen gespalten und die Spaltungsprodukte in einem Agarosegel aufgetrennt. Nach der Trocknung der Gele und deren Exposition konnte die Reihenfolge der Spaltstellen leicht abgelesen werden, da nur Partialfragmente mit der markierten Eco-RI Stelle erkennbar waren. Auf diese Weise wurden die Restriktionskarten von Hpa II, Pst I, Hae III und Hha I bestimmt, die in Abb. 33 ange-

geben sind. Die spätere Sequenzierung bestätigte die Lage der meisten Spaltstellen und erlaubte die Ergänzung zusätzlicher, mit dieser Methode nicht eindeutig kartierbarer Sequenzen.

Bereits aus der Restriktionskarte lassen sich verschiedene auffällige Eigenschaften des Klons pSt 608 erkennen. Eine ungewöhnliche Häufung vieler Spaltstellen von Hpa II (CCGG), Hha I (GCGC) und Hae III (GGCC) findet sich in der Umgebung der Sma I-Stelle (CCCGGG). Alle diese Enzyme spalten Sequenzen, die ausschließlich CG-Paare enthalten. Die Sequenzierung bestätigte später die daraus abgeleitete Vermutung, daß dieser Bereich einen extrem hohen GC-Gehalt aufweist. Außerhalb dieser Region fanden sich nur sehr wenige Hpa II Stellen, bei deren Spaltung zwei Fragmente von 1.9 und 1.7 kb Länge auftraten. Diese Fragmente entsprachen den durch die Hybridisierung genomischer Spaltungen bestimmten Fragmenten, die unmethylierte Hpa II-Enden aufwiesen (s.C.3.7.). Die drei Hpa II-Stellen, die den nichtmethylierten Sequenzen des Genoms entsprechen sind in Abb. 33 angegeben.

Eine weitere Besonderheit zeigte das Muster der Hae III Spaltung, das auf die Repetition eines kurzen Sequenzbereichs von etwa 100 bp Länge hindeutete. Die Hae III-Stellen folgen in Abständen von 100-100-100-200-200-200-400-400 bp. Ob es sich dabei tatsächlich um eine Folge sich wiederholender Einheiten handelt, ist noch unklar, da dieser Bereich noch nicht sequenziert wurde. Falls es sich aber um eine repetitive Sequenz handeln sollte, dürfte diese im Genom nur an dieser Stelle vorhanden sein, da die Hybridisierung genomischer Spaltungen mit der klonierten Region des p550-Gens keinen Hinweis auf eine im Genom häufig auftretende Sequenz ergab (s.C.3.6.).

Diese Besonderheiten (Anordnung und Methylierung der Hpa II Stellen: extrem hoher GC-Gehalt in Teilbereichen; mögliche Repetition) im Zusammenhang mit der Lage am 5'-Ende des Gens könnten auf eine mögliche Bedeutung als regulatorische Elemente hindeuten.

6.2 Grobe Lokalisierung der kodierenden Bereiche

Die Hybridisierung von Teilfragmenten des Klons pSt 608 wurde bereits früher genau beschrieben (s.C.4.2.). Der umgekehrte Weg, die Hybridisierung markierter RNA mit Fragmenten des Subklons, sollte es erlauben, die mRNA Sequenzen bestimmten Bereichen zuzuordnen.

Da die mRNA des p550-Gens nur in geringen Mengen im Gemisch aller RNAs vorliegt, mußte sie zuerst angereichert werden. Dazu wurde die RNA aus Sterna (17 Tage alte Embryonen) in einem denaturierenden (5M Harnstoff) Agarosegel aufgetrennt (105) und daraus die RNA im Bereich von 4.3 und 6.5 kb isoliert. Durch Behandlung dieser RNA mit Alkali (50 mM Tris, pH 9.5) ließen sich Bruchstücke mit einer Länge von 50-150 bp Länge erzeugen, die danach mit Polynukleotidkinase und $\gamma^{32}\text{P}$ -ATP an den 5'-Enden markiert wurden (106). Diese Mischung von markierten Bruchstücken konnte dann zur Hybridisierung eingesetzt werden.

Die hybridisierenden Fragmente von pSt 608 sind in Abb. 33 angegeben. Besonders deutliche Hybridisierung fand sich mit 1.95 kb/ Hpa II, 1.18 und 1.5 kb/ Pst I und 2.15 kb/Hha I. Schwächer reagierten die Fragmente 1.7 kb/ Hpa II und 1.15 kb/ Pst I. Damit scheint der größte Teil des Subklons kodierende Sequenzen zu enthalten. Eine genauere Lokalisierung von Exons war mit dieser Technik nicht möglich.

6.3 Sequenzierung von Teilen des Klons pSt 608

Die drei Subklone. 608A mp8, 608A mp9 und 608B mp8 (Abb. 28) ließen sich dazu verwenden, die Sequenzen der Umgebung der Bgl II und Hind III Stelle zu bestimmen. Mit Hilfe der Didesoxymethode wurden jeweils etwa 300 bp der Randsequenzen des jeweiligen Inserts bestimmt.

Konstruktion von Deletionsmutanten des Klons 608A mp9

Zur Sequenzierung des 3.7 kb langen Bgl II/Hind III-Fragments wurde eine Methode verwendet, die es erlaubte, auf relativ einfache Weise die Sequenz dieses großen, in einem M13 Vektor klonierten (608A mp9) Fragments in einer definierten Richtung zu bestimmen. Dazu werden Deletionen unterschiedlicher Länge im Bereich des Inserts erzeugt. Um eine gleichmäßige Verteilung von Deletionen zu erreichen, wurden zwei Verfahren beschrieben: entweder verwendet man die Exonuklease Bal 31 (107), oder es wird der Effekt ausgenutzt, daß DNase I in Anwesenheit von Mn^{++} -Ionen zu Doppelstrangbrüchen führt. Mit der zweiten Methode (108), die auf elegante Weise die Eigenschaften des M13 Systems ausnutzt, wurden Deletionsmutanten von 608A mp9 hergestellt. Dieses Verfahren bietet gegenüber der 'Shot-Gun' Klonierungs- und Sequenzierungsmethode (65) den Vorteil, daß alle Sequenzen nur von einem Strang erhalten werden und deshalb einfach zu einer Gesamtsequenz zusammengefügt werden können.

Das Prinzip der Konstruktion von Deletionsmutanten ist in Abb. 34 schematisch dargestellt. Der Klon 608A mp9 enthält das Bgl II/Hind III-Fragment so in die Bam HI- und Hind III-Stellen des Vektors eingebaut, daß die Leserichtung der Sequenzen der Orientierung der mRNA Sequenz entspricht: 5'-Bgl II-Hind III-3' (s.C.4.2.). Die doppelsträngige, replikative Form des Klons wird durch DNase I in Anwesenheit von Mn^{++} -Ionen, einmal gespalten und die entstehenden linearen Moleküle aus dem Gel isoliert. Im nächsten Schritt wird mit einem Restriktionsenzym gespalten, dessen Spaltstelle zwischen der Bindungsstelle des Sequenzierungsprimers und dem Beginn des Inserts liegt. Im Falle des Klons 608A mp9 ist dafür die Eco RI Stelle des Polylinkers besonders geeignet. Nachdem die Einzelstrangenden der entstehenden Fragmente mit DNA Polymerase I (Klenow-Fragment) aufgefüllt worden sind, können sie unter Bedingungen ligiert werden, die zu einer Zyklisierung führen.

Die Selektion der entstehenden Deletionsmutanten bei der

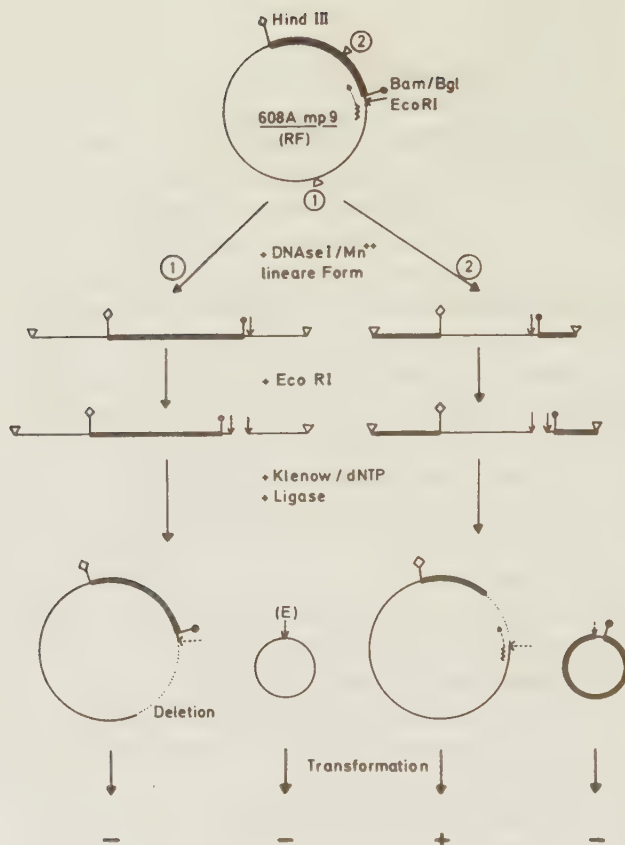


Abb. 34 : Prinzip der Herstellung von Deletionsmutanten des Klons 608A mp9 (Beschreibung im Text)

nachfolgenden Transformation läßt sich verstehen, wenn man die beiden unterschiedlichen Möglichkeiten der DNase I Spaltung betrachtet: spaltet die DNase I im Vektoranteil des Klons (Fall 1), so entstehen nach der Spaltung mit Eco RI und der Ligation zwei Moleküle, die keine vollständigen Replikons mehr repräsentieren und deshalb nach der Transformation nicht mehr zur Ausbildung von Plaques führen können; erfolgt die Spaltung mit DNase I jedoch im Bereich des klonierten Inserts (Fall 2), so bleibt bei den folgenden

Schritten der Vektoranteil intakt, und es geht nur der Teil des Inserts verloren, der zwischen der DNase- und der Eco RI Spaltstelle liegt. Nur in diesem Fall können Plaques entstehen. Die Phagen dieser Klone unterscheiden sich alle in der Länge des deletierten Bereichs zwischen Primerbindungsstelle und der DNase I Spaltstelle. Bei der Sequenzierung erfaßt man deshalb verschiedene Bereiche innerhalb des klonierten Fragments.

Die Bedingungen der Spaltung mit DNase I wurden in einem Vorversuch so ausgetestet, daß die isolierte replikative Form von 608A mp9 jeweils nur an einer Stelle der DNA gespalten wurde: 2.5 µg DNase I/ µg DNA/ 1 mM $MnCl_2$ / 20°C. Die zirkuläre, doppelsträngige Form geht dabei in die lineare Form über, was durch Gelelektrophorese leicht verfolgt werden konnte. Die entstandenen linearen Moleküle mit einer Länge von 10.8 kb wurden aus dem Gel isoliert und mit Eco RI gespalten. Alle entstandenen Einzelstrangenden der Eco RI- und der DNase I-Spaltung wurden danach mit Hilfe von DNA-Polymerase (Klenow) und allen vier Desoxynukleotiden zu 'Blunt Ends' aufgefüllt. Diese Moleküle konnten nun durch Ligation wieder zirkularisiert und in JM 103 transformiert werden.

Die Ausbeute an Plaques war gering: nur etwa 45 Plaques/ µg linearisierter DNA. Zwei Gründe sind dafür verantwortlich: erstens spaltet DNase I 2/3 der Ausgangs-DNA (Vektor 7.3 kb/ Gesamtklon 10.8 kb) im Vektoranteil (Fall 2 in Abb. 34) und zweitens erfolgte die Zirkularisierung der Fragmente nur mit einer geringen Effizienz.

Bestimmung der Insertlängen der Deletionsmutanten

Alle Klone wurden in kleinem Maßstab kultiviert, die Phagen-DNA isoliert und deren Größe durch Elektrophorese in Agarosegelen abgeschätzt. Alle Klone wiesen unterschiedliche Längen auf, die zwischen denen des Ausgangsklons 608A mp9 und des Vektors M13 mp9 lagen. Diese Abschätzung gibt jedoch

nur die Größenordnung der einzelnen Klone an. Eine genauere Bestimmung der Länge des restlichen Inserts ist möglich, wenn man die Komplementarität der klonierten Inserts in 608A mp9 (bzw. der davon abgeleiteten Deletionsmutanten) und 608A mp8 ausnutzt. Hybridisiert man die einzelsträngige DNA dieser Klone, bildet sich ein Doppelstrang im Bereich der komplementären Sequenzen. Baut man die verbliebene einzelsträngige DNA durch S1-Nuklease ab, ist nur der hybridisierte Bereich geschützt.

Abb. 35 zeigt die Auftrennung einer Reihe solcher S1-Abbauprodukte. In der Tabelle 12 sind die Insertlängen der Klone angegeben, die auf diese Weise bestimmt wurden. Nur ein Klon (608A-3) war kleiner als der Ausgangsvektor M13 mp9. Ein solches Ereignis läßt sich dadurch erklären, daß die DNase I-Spaltung im Vektoranteil - in unmittelbarer Nachbarschaft des Inserts - erfolgte, aber keine essentiellen Teile des Vektors verloren wurden, die das Wachstum des Phagen beeinträchtigen.

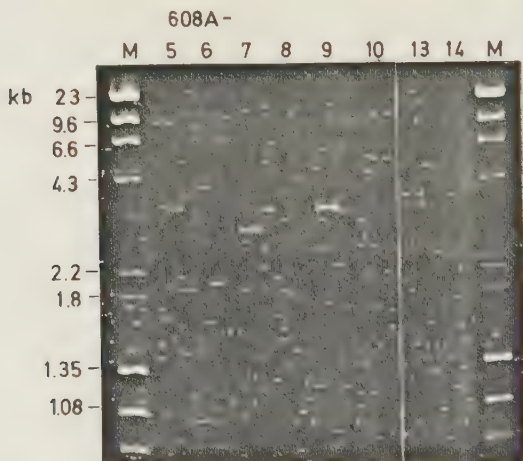


Abb. 35 : Bestimmung der Insertlängen verschiedener Deletionsmutanten

Je 1 μ g Phagen-DNA der angegebenen Klone wurde mit 1 μ g 608A mp8 hybridisiert und mit S1-Nuklease abgebaut (0.8% Agarosegel)

Tabelle 12 : Durch S1-Kartierung bestimmte Insertlängen
eines Teils der Deletionsmutanten

| Klon | Insert (kb) | Klon | Insert (kb) |
|--------|-------------|---------|-------------|
| 608A | 3.7 | 608A-13 | 3.50 |
| | | -14 | 2.40 |
| 608A-1 | 0.95 | -23 | 2.95 |
| -2 | 1.45 | -25 | 1.90 |
| -3 | (0) | -26 | 0.20 |
| -4 | 1.18 | -28 | 2.60 |
| -5 | 1.95 | -29 | 0.35 |
| -6 | 2.90 | -32 | 2.60 |
| -7 | 1.55 | -33 | 2.40 |
| -8 | 3.40 | -35 | 2.30 |
| -9 | 1.70 | -39 | 0.40 |
| -12 | 0.90 | -40 | 3.35 |
| -13 | 3.50 | -41 | 2.10 |

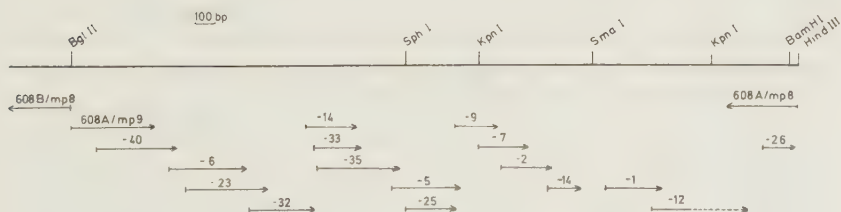


Abb. 36 : Länge und Orientierung der gelesenen Einzelsequenzen

--- : Keine Unterscheidung G oder C

-NN : Deletionsmutanten : 608A-NN

Sequenzierung der Deletionsmutanten von 608A mp⁹

Die Sequenzen von 16 dieser Deletionsmutanten wurden meist mehrmals mit Hilfe der Didesoxymethode bestimmt, wobei jeweils 200 bis 400 bp gelesen werden konnten. Ein Beispiel einer solchen Sequenzierung ist in der Abb. 37C gezeigt. Durch die vorhergehende Bestimmung der Insertlängen war es leicht, die Einzelsequenzen zu einer Gesamtsequenz zusammenzufügen. Lage und Länge der einzelnen sequenzierten Bereiche sind in Abb. 36 angegeben. Es läßt sich erkennen, daß ein Großteil der Sequenzen bei der Sequenzierung von mindestens zwei unabhängigen Klonen erhalten wurde. Die gelesenen Sequenzen konnten auf diese Weise untereinander verglichen und überprüft werden.

Sequenzierung der GC-reichen Bereiche

Anfangs wurden zur Auftrennung der Fragmente der Sequenzreaktionen Gele von 80 cm Länge und 0.4 mm Dicke verwendet (108). Die Temperatur des Gels erreichte allerdings, bedingt durch die Länge und damit den schnellen Wärmeaustausch mit der Umgebung, nicht zu allen Zeiten die zur völligen Denaturierung der DNA-Fragmente notwendige Temperatur von 50-55°C (64). Auch die Verwendung von Puffern mit höherer Ionenstärke (1.5-2 x TEB) brachte keine entscheidende Verbesserung, obwohl die Temperatur der Gele besser kontrollierbar wurde. Dies hatte bei den meisten Klonen keine sichtbaren Auswirkungen, die Gele waren eindeutig lesbar. Die Sequenzen der Klone 608A-1, -2, -4 und -12, waren unter diesen Bedingungen jedoch nicht lesbar. Ein Beispiel ist in der Abb. 36A gezeigt.

Das Auftreten von Banden, die den gleichmäßigen Abstand der einzelnen Fragmentlängen storten und zusätzliche Basen vortauschten, machten es unmöglich eine eindeutige Sequenz zu lesen. Offensichtlich reichten die denaturierenden Bedingungen im Gel (5 M Harnstoff, etwa 50°C) nicht aus, um

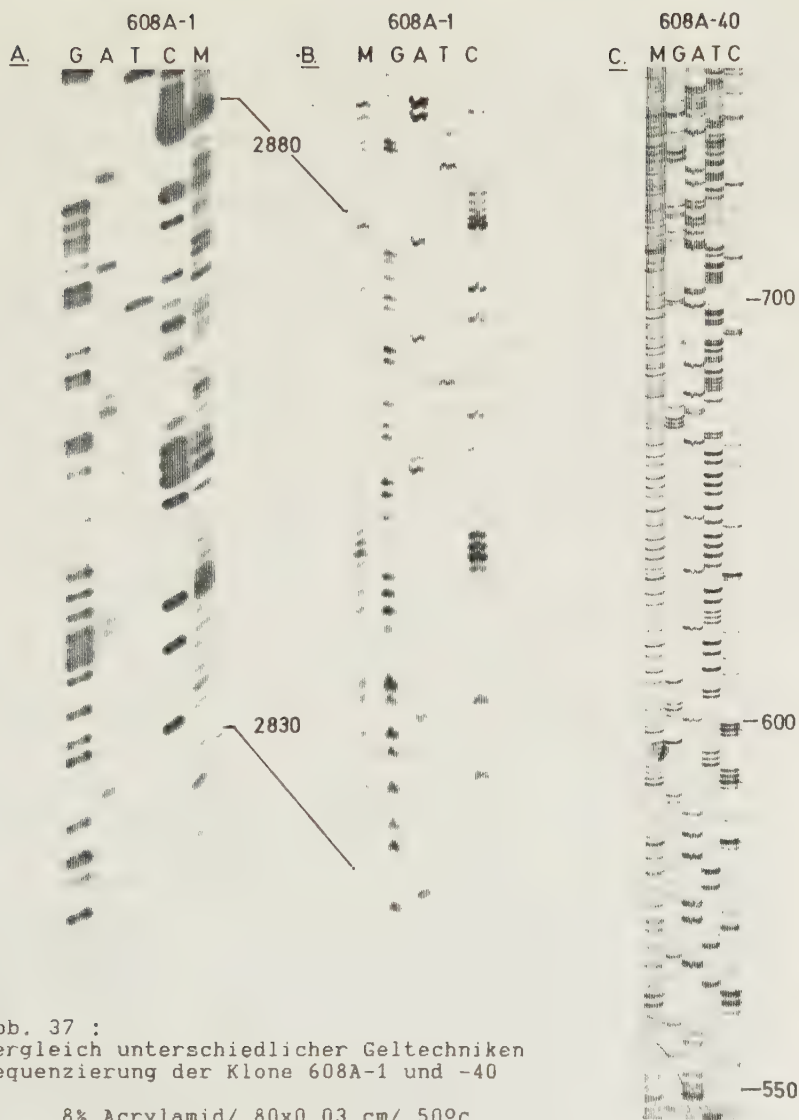


Abb. 37 :
Vergleich unterschiedlicher Geltechniken
Sequenzierung der Klone 608A-1 und -40

A. 8% Acrylamid/ 80x0.03 cm/ 50°C

Verwendung von $\alpha^{32}\text{P}$ -dATP

B, C. 6% Acrylamid/ 60x0.02 cm/ 65°C

Verwendung von $\alpha^{35}\text{S}$ -dATP

Positionsangaben : s. Abb. 37

M : Mischung der Ansätze G+A+T+C

alle Sekundärstrukturen der Fragmente zu eliminieren. Alle Klone, die dieses Verhalten zeigten, lagen in der GC-reichen Region von pSt 608. Eventuelle Sekundärstrukturen der DNA, die - bedingt durch den hohen GC-Gehalt - besonders stabil sind, mußten deshalb durch drastischere Bedingungen bei der Elektrophorese aufgelöst werden.

Eine Verbesserung der Lesbarkeit der Gele von Sequenzen dieses Bereichs war erst möglich durch die Verwendung sehr dünner Gele (0,2 mm), deren Temperatur durch zusätzliche Thermostatisierung konstant bei 65°C gehalten werden konnte (109). Die Gele wurden dazu über ein bifunktionelles Silanderivat kovalent an eine Glasplatte gebunden. Die Thermostatisierung erfolgte über die zweite Platte, durch die Wasser mit konstanter Temperatur gepumpt werden konnte (LKB-Macromould). Die Gele wurden nach dem Lauf getrocknet und ergaben deshalb scharfere Bandenmuster. Durch die Thermostatisierung der Gele, die zudem das 'Smiling' einzelner Bahnen verhinderte, und die Verwendung dünner Gele konnten die bisher nicht lesbaren Sequenzen bestimmt werden, wie der Vergleich in Abb. 36 zeigt.

Ein weiteres Problem bei der Sequenzierung dieser Klone war häufig die geringe Länge der entstehenden Fragmente bei den G und C spezifischen Reaktionen. Der hohe GC-Gehalt, führte bei dem verwendeten ddNTP/dNTP Verhältnis sehr schnell zu Kettenabbrüchen. In diesen Fällen wurde deshalb anstatt eines ddNTP/dNTP Verhältnisses von 40:1, nur mehr ein Verhältnis von 15:1 oder 30:1 verwendet. Die Verteilung der Fragmentlängen konnte damit deutlich zu größeren Fragmenten hin verschoben werden.

Neben diesen Verbesserungen der Lesbarkeit GC-reicher Sequenzen durch Thermostatisierung von Gelen war es möglich, durch die Verwendung von $\alpha^{35}\text{S}$ -dATP als markiertem Nukleotid (anstatt des ^{32}P -Analogons) die Auflösung der Banden erheblich zu steigern (111). Die schwächere Strahlung des Schwefelisotops führte zwar im Vergleich mit dem ^{32}P -Isotop zu einem bedeutend schwächeren Signal, die Scharfe der Banden und deren Unterscheidbarkeit, besonders im Bereich größer

Fragmente, nahmen jedoch deutlich zu (Beispiel : Abb. 36C). Durch Trocknung der Gele nach dem Lauf konnte die Expositionszeit dennoch kurz gehalten werden (20-50 Stunden). Bei der Verwendung von $\alpha^{35}\text{S}$ -dATP mußte die Inkubationszeit der Ansätze etwas verlängert werden (20 Min.), um die geringere Affinität des Klenow Enzyms für dieses Nukleotid auszugleichen.

Durch die Anwendung aller Verbesserungen (Thermostatisierung, dünne Gele, $\alpha^{35}\text{S}$ -dATP) war es möglich, bei Gelen von 60 cm Länge meist mehr als 350 Basen sicher zu lesen, wobei teilweise noch Sequenzen von mehr als 450 Basen lesbar waren.

Zusammenfassung der Sequenz von pSt 608

Auf die beschriebene Weise ließen sich, ausgehend von der Bgl II Stelle, 2650 Basen eindeutig bestimmen. Der folgende Bereich von etwa 150 Basen in der Umgebung der Sma I Spaltstelle, konnte nicht fehlerfrei gelesen werden. Die Sequenz dieses fehlenden Bereichs wurde nach der Methode von Maxam & Gilbert bestimmt, und wurde in Abb. 38 mit verwendet (75).

Abb. 38 : Teilsequenz des Klonen pSt 608.

AGATCT : Bgl II Restriktionsstelle

ZZ : G oder C (S.G.2.12.)

Sequenzfolge 2650-2780 aus (75)

| | | | | | |
|-----|---------------------|------------|------------|------------|------------|
| 1 | TGCTGGCAGT | ATCTCTTGGG | GGGGGGGGGG | GGCCGCTGGG | GGAGAAGTTG |
| 51 | CCACGTGTTT | ACTTTGTATT | GTGATTGCCG | AGGGAAGGAG | AGCAGGGACA |
| 101 | CTGAAAAGGT | AAACAAAAGA | GCTCCAGCCC | AATGGCCGAA | GCAGAGTAAA |
| 151 | GTAAAAAGTTT | CCTTTTGAAG | AATAGGAAGG | AAATATGCCT | GCTTTGTTCT |
| 201 | GCAGGATCTC | TGCAATCCCC | ACGTCTTGGC | TAACCGTGGT | ATTTAACCAC |
| 251 | AAGTGTCCAA | ACAAGCTCTG | GGAGGCTGGC | TCCGGGTAAT | TATTGGTGGA |
| 301 | AGATCT TTATT | TATTTTTTCT | AGTGCAGGCA | GAACAGATAG | TGCTGCAGCA |
| 351 | CTGCTCCAAA | GGCACCTGTG | TAAGATGAGC | ACACAGTGTG | GTAATTATGT |
| 401 | AAATGAAAAG | CACATTCCGA | TCAGCCTAAA | CTACCCAAAA | TTGGGATGAA |
| 451 | AGATGGTCAG | AAAAGAAACC | TTTTTCACTA | GTGGGGGAGT | CAGGCATGGA |
| 501 | ATGGCTCACC | CAGGAGCTTG | GTGCAGCTTT | GCAGGGTCCA | AGAGGTGATG |
| 551 | AAACGAAGCC | CTGAGTGCAC | AGTGTGACCA | GCAGGGCCCC | TTTGGCCCGG |
| 601 | GGTTGGCTGC | TCTGCATTTT | TGTGCCATCT | CTCTCCTACT | GCTCTGTGCT |
| 651 | GTGCATTGGG | GAAAGCATTT | TCTGATTCTT | TTCCCGTTCT | TCAGGCAAGT |
| 701 | TTTGTTCCTG | TAAAGCAGCA | CAGGCAGATG | CTCCATTTCA | CTTGGTGGTG |
| 751 | TCGTGAGTCT | CCCAGGCAGA | AGACCAGAAC | CCAGAACTTT | GCCTCCCTTT |
| 801 | CAAACCTTGC | AGTTCTTCAC | TTTCCCTATG | GTCTCTCTTG | AGGTGCGGGT |
| 851 | AGGAACCTCA | CGTGACTGGT | GCAGCGTGTG | TTATTTCTTT | GTGAAATTCC |
| 901 | CAGACTTCTC | TCCCAATGCA | AGTGGGAGGA | CTACCTCTGA | CCTTTCTCTT |
| 951 | GCTGTTGGAT | TGATTTTTAT | TCAGCTGTCC | ATTGACTTTC | TTCCTCATCT |

1001 CTGAAATGGC ATTTGCTTTG CGATAGTACC TGGCATCCAA ACAATTGGTA
 1051 CGGAATACTG AATCCACGGA GCAGATCAGA CCTTGTGTCC ACAGCTGAGA
 1101 GAATCCCACC AGTCAGTGCA GTATCAGTGC AGTAAGCTGG CAGCAGAACC
 1151 AGGAGTGGGG CTCAGTCTCT GGACTGGTCA GTGTTAACCA CTGGGTTGGG
 1201 TGGCCTCGCA GTATTCAGT GGCTTGCTTT GCCTGCAATG TGTGGCTTTT
 1251 TCTCTCCCTG CGTTTACAGA GCTTGAAGCT ACTTGTCCAT GTTAGAAGTC
 1301 CATTGCCACC AGTAAAGTTG TCCTGTTCTT ATGCACTTTG CTTGCTCCCC
 1351 TTGTGTAGAG CGCCCTGTGA CAGCCTCAGA AGGGCAGCGT ATATGTTGTG
 1401 TGCTGGCTGA CGATTGCATA ACTTTATTGT CAGAAATGCC CTGAGTAAAT
 1451 AAATTTGTCT TTTGAACGAG CATTTTGGCC TGAGAGGCCT GCAGAGGTTA
 1501 GTCACCTTGC TTCCGTATCT GCTCAGCTCT TCTGCTTGTG CAAGAGGAGA
 1551 GCTGCCCTCT ATAAACAACG CATTTACCAG GAGAAAGGCA AGTGGTGCTT
 1601 GTGCTTCCCT GGGACCCGTA TTTATAGCAC CTGGGTTACC AGAGACAGAC
 1651 CACCTGGCTG CGTGCCCTCC TGCCACCCCC TCCCCAGTCA CTGCACATGA
 1701 CTGCAGCTGG GGTACTAGG GGGCTCAGAG GGGCTGCGAT TTCCCTGGCC
 1751 TTCTGAAGGT GGTCAAGATT CACTGGGCTC AGCTGAGGGC AGAATAGTTT
 1801 GGGCTGGTAA TGATGTGTGA GTGAGGAGAA GAGGAGTGGG TGGGTGCTTT
 1851 AGAAGTTGCA ATAAGAAAAG AGAATCTGGT CTGAAAGCCA CCAGGATGCA
 1901 TGCGGCATGG TCACACTAG TGCCAGGCAG TTCTTAGGGC GTTGTTTTCT
 1951 GTATGCTCTT CGTAGAGGT TGAGAAACAC AGAAGGATCT CCATAATTCT
 2001 ATCCTTAACT TCAGTGTAAT AAAGGTTTTT CAAACTGTGC TACTGGGAGA
 2051 TGGCACGGCA GGCTGTAGTG ATGGCTTCAG GGCGGCATCT CCTGCTCTCT
 2101 CTGGATGGGA ACCCCAAATG GGCTCCGGCA TGAGTGGCCG AGTCTCTCTT
 2151 TACTGTGCTT TTTAATGATG CCGAGGACTT CTTTATCGTG CTGCAGCCGA
 2201 GGGATGTGGG TGTTTTGGGG CGGTCCCTCG CGTGCGCTGT GACTZGGCTG
 2251 TATGGGGTGG GTACCGCAGC ACAACGGGGC GCTGATGTGA GCGTTTTAGC
 2301 AACGGGGCCG TCCTGAAACA GGAAATCGAG TCGCGGCCCA GGGGAGCGCG
 2351 GTCGGGTTGC CTTCCAGAGA AGTTTCTGGC GAGGAAAATA GGGCTCTCCT
 2401 CTCGAGACGG CGTGCAAGAA GAGGCCGGAA GAATCCCCCG ACGAGCTGCG
 2451 GTCGCGCCTG CGGCAGCTCG GCCGCGCTGT TGTGCGCCGT CGCGTGCCTG
 2501 CCGTCTGTGCC CCGCTCCGCT CGCGCGCCCC GCGGGGAAGG GGCACGTCCC
 2551 GCTGACGTGC GCGCGCGGGA TTGCAGGCAG CGGCCCTGCC TATATAGGGC
 2601 CGCGGGAGCG TCGCGCCGCT CTCCCAGGAG CGCTCCCGGT TCTCCAGCGA
 2651 GCGCGGCGAG CGCGCGCGCG GCCGGAGCCG CGCGCGCGCG GGGCGCGGGG
 2701 CGCGGAGGGG AGCGGAGCGG GCGGAGGGGG GCQCGCGACC SCCCAGGGGA
 2751 AGCTCCGGCG GCCGCGGAGA CGAGCGGGGC GAGGAACGCG CGGGGGGCGC
 2801 CGGGGAGGAG AGGAGAGGAG GAGAACGGAG CGGAGCGGAG GGGCCCCGCG
 2851 GAACGGCGCG CGGAGCGGCC GCGGACCCCC CGTGCGGTGA CAGCGGCGCG
 2901 GCCTGGCGCT GCGCGCGCCC CGCCCCGGT GCGGCGAGTG CTGAGGTGGC
 2951 GCTGGCGGCG CZZCCGGGCC TGGGCCCGZC GGCTZCAZZZ ZZZZZZZZZZ
 3001 ZZZAZZZTZZ ZZZZZZZZZZ ZZZZZZZZZZ AZZAAZZZZZ ZZZGGGAGG
 3051 AAGGGGCCCG GCCGCTGGGC GCCCGAACGG AGCGCGGGCC GCGGGTGGC
 3101 CGGGCACAAA GTGAGCGCTG TGTGAGCACA GGAACCCGAG GTAGGGGAGG
 3151 TGGCGGCGAG CGGGTTAGGA GGAGAGAAAA CAGGACAZA GCTCGGATTG
 3201 AAGGCTGGCC CGCTCCAGCA GCGCGCGGTC AGTTGCGCG GTZZCGGAGG
 3251 AAGCGGTZGA AGGGCTCATG CTGCCAZZZA ZZZAZZZZZZ ZZZZZZZAZ
 3301 ZAZZZAAZZZ ZZZZZZZZZZ ZZZTAZZTZZ ZZZZZZZTAZ ZAAZZAZZTA
 3351 ZZZZZAZAZZ ZZAAZAZZAZ ZTTZZZEZTA ZAAZZTZZT AZZZAZAZZ
 3401 AZZZAZZZZZ ZZZTTZZTAA CCTTAZTZZT ZTGCAGCTGT AGGCACCGTA
 3451 ACACCAAACG CAGATATCTC GGTAAAATAT TGCCCCGGCTG CCTGGTTCTG
 3501 TCCACCAGCA TTGGCTGTGT TGGCAGTGCG GGGCAAGACT TTGCTTTTGT
 3551 TTCCCTTTCT ACAACCGTTT CCGTATGGTT GGGGCTGGCC CTCTGATAGGA
 3601 GCTACTCTCG GCACGATTTT ATTCTCCCTA TTTGAGCGTG CTCTGAACTT
 3651 GCAAAGCTCG CTGTGGCTGT GCTGCAGCCG GCCGTAAGCC AGGCGGTTCT
 3701 GATCCTGGTG ACCCAAGCGG CCGTGTGCCG GCCTGCCGTC GCTGTAAGCT
 3751 T

Die restlichen 950 Basen bis zum Ende des klonierten Inserts, der Hind III Stelle, konnten jedoch wieder eindeutig sequenziert werden. In zwei kurzen Bereichen war es zwar möglich, den Raster der Sequenz eindeutig zu verfolgen, eine Unterscheidung der Basen G und C war jedoch aufgrund der Probleme bei der Sequenzierung der GC-reichen Region nicht sicher möglich. Diese Sequenzen finden sich an den Positionen 2990-3040 und 3280-3420.

Damit ergab sich eine Gesamtsequenz von 3750 bp. Die Nummerierung der Basen in Abb. 38 beginnt willkürlich mit der am weitesten 5'-terminal gelegenen Base der Sequenz von 608B mp8. Die Bgl II Stelle findet sich dabei an Position 301.

6.4 Analyse der Sequenz von pSt 608

Ein Vergleich der Sequenz dieses Teilstücks von pSt 608 mit der vorher gefundenen Restriktionskarte bestätigte die Lage der meisten Spaltstellen. Zwei sehr nahe bei bereits kartierten Stellen liegende Pst I Stellen konnten erst durch Sequenzierung entdeckt werden. Im Falle der Restriktionsenzyme Hpa II, Hae III und Hha I war erst durch die Sequenzierung die genaue Angabe der Spaltstelle, besonders im Bereich der GC-reichen Region, möglich. Diese Daten wurden teilweise bereits in Abb. 33 mit verwendet.

Verteilung des GC-Gehalts innerhalb der Sequenz

Bisher wurde nur pauschal von einem 'GC-reichen Bereich' gesprochen. Die genaue Verteilung ließ sich erst dadurch darstellen, daß der jeweilige GC-Gehalt eines Teilfragments von definierter Länge bestimmt, und mit diesem 'Fenster' die gesamte Sequenz durchsucht wurde. Als Testlänge wurden 100 Basen festgelegt, um sehr kurze, lokal begrenzte GC-reiche Sequenzen auszuschließen. Das Ergebnis ist in Abb. 39 gezeigt. Die aus der Restriktionskarte vorhergesagte GC-rei-

che Region in der Umgebung der Sma I-Spaltstelle ist sehr deutlich zu erkennen. Grob eingegrenzt wird diese Region von den beiden Kpn I-Spaltstellen (Pos. 2260, 3347). Der GC-Gehalt liegt in diesem Bereich immer mindestens bei 60%. Mehr als 70% GC findet sich zwischen den Positionen 2400-3100 und 3200-3300. Zwei kurze Bereiche um die Positionen 2700 und 2900 weisen sogar Werte von mehr als 90% GC auf. In einer Region von mehr als 700 bp Länge fand sich somit in der Sequenz von pSt 608 ein extrem hoher GC-Gehalt von mehr als 70%.

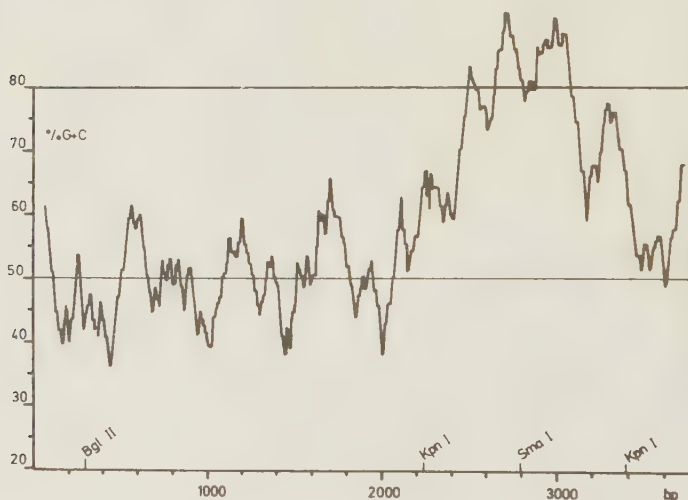


Abb. 39 : GC-Gehalt der Teilsequenz des Klon's pSt 608
(Positionsangaben s. Abb. 38)

Die GC-Verteilung in anderen Bereichen der Sequenz weist ein anscheinend regelmäßiges Muster von erhöhtem GC- und AT-Gehalt auf. So tritt ein GC-Gehalt von jeweils etwa 60% in der Umgebung der Positionen 600, 1200, 1700, 2200 und 3200 auf.

Sucht man die gesamte Sequenz mit geeigneten Computerpro-

grammen auf die Existenz von 'Inverted Repeats' ab, von Sequenzen, die potentiell zur Ausbildung von Sekundärstrukturen führen könnten, so finden sich solche in großer Zahl in der Region des hohen GC-Gehalts. Ob die denkbare Ausbildung stabiler 'Hairpin'-Strukturen in diesem Bereich eine funktionelle Bedeutung für die Expression des p550-Gens oder das Processing der mRNA hat, läßt sich noch nicht entscheiden.

Anordnung der 5'-CG Sequenzen

Innerhalb des GC-reichen Bereichs finden sich extrem viele potentiell methylierbare 5'-CG Sequenzen, sichtbar auch an der Häufung der Hpa II- und Hha I-Stellen. In den übrigen Bereichen sind diese Sequenzen deutlich unterrepräsentiert. Für drei dieser Hpa II-Stellen (eine Stelle nicht sequenziert/ Positionen : 282, 2125) konnte gezeigt werden, daß nur diese Sequenzen im Genom nicht methyliert vorliegen, während im gesamten restlichen Gen alle erkennbaren CCGG-Sequenzen methyliert vorliegen. In welchem Zustand die große Zahl von CG-Sequenzen des GC-reichen Bereichs vorliegen läßt sich mit Hilfe der angewandten Methoden nicht nachweisen.

Suche nach möglichen Promoterstrukturen

Auf Grund fehlender Daten über den Startpunkt der Transkription und die Lage und Länge der Exons ist eine Suche nach solchen Strukturen nur möglich durch den Vergleich mit bekannten Prototypsequenzen. Die dabei gefundene, hypothetische Struktur des Gens kann durch gezielte zusätzliche Versuche später überprüft werden.

Der Vergleich der Promoterregionen einer Vielzahl von Genen führte zur Aufdeckung von zwei typischen Sequenzfolgen, die eine wesentliche Rolle bei der Transkription durch die RNA-Polymerase spielen. Die konservierte 'TATA-Box' findet sich etwa 25 bp vor dem Startpunkt der Transkription.

In einem Bereich von etwa 70-80 bp vor dem Startpunkt findet sich häufig die 'CAT-Box' (104). Die gesamte, bisher bestimmte Sequenz von pSt 608 wurde nach Ähnlichkeiten mit diesen beiden Sequenzen abgesucht. Dabei fanden sich keine absolut homologen Sequenzen; einige Bereiche aber wiesen eine recht gute Übereinstimmung mit den Prototypsequenzen auf, besonders die Sequenzen (A,B,C in Tab. 13) in der Umgebung der Bgl II-Stelle (Pos. 300).

Tabelle 13 : Promoterähnliche Sequenzen in pSt 608

Positionsangaben : s. Abb. 38

NN : Sequenz identisch mit Prototypsequenz

| | | | |
|------------|-----------------------------|-----------|------|
| Prototyp : | GGCCAATCT-----TATATAT-----A | | |
| | T | A A | |
| | (-70/-80) | (-26/-34) | (+1) |
| A : | AGGAAATAT-----TATTTAA-----A | | |
| | 180(-82) | 240(-23) | 263 |
| B : | GGCTAACCG-----TAATTAT-----A | | |
| | 228(-84) | 287(-25) | 312 |
| C : | AAACAAGCT-----TATTTAT-----A | | |
| | 259(-77) | 307(-30) | 337 |
| D : | ? -----TATATGT-----A | | |
| | | 1390(-30) | 1420 |
| E : | ? -----TATAAAC-----A | | |
| | | 1560(-31) | 1591 |

Über eine eventuelle Bedeutung einer dieser Stellen kann bisher noch keine Aussage gemacht werden. Durch Kartierung des 5'-Endes des p550-Gens sollten die tatsächlich verwendeten Promoterstrukturen gefunden werden können.

Hypothetisches Exon-Intron Muster des Klons pSt 608

In früheren Versuchen konnte gezeigt werden, daß die kodierenden Sequenzen über den gesamten Bereich des Klons pSt 608 verteilt vorliegen (s.C.6.2.). Versuche, die Lage und Längen der Exons durch R-Loop Kartierung und S1-Experimente zu bestimmen, führten bisher noch zu keinem eindeutigen Ergebnis. Ursache dafür ist möglicherweise die schnelle Renaturisierbarkeit der GC-reichen Sequenzen und die Ausbildung von Sekundärstrukturen in diesem Bereich. Aus diesem Grund wurde versucht, ein hypothetisches Exon-Intron Muster aus der Sequenz abzuleiten.

Bei einer Vielzahl von Genen konnten die Exon-Intron Übergänge sequenziert werden und daraus typische Sequenzen abgeleitet werden (103). Diese Prototypsequenzen wurden bereits bei der Suche nach entsprechenden Stellen des p550-kodierenden Bereichs verwendet (s.C.5.2.). Die Sequenz des Klons pSt 608 wurde nach ähnlichen Sequenzen abgesucht, wobei sich eine große Anzahl von homologen Sequenzen finden ließ.

Um jedoch die Lage hypothetischer Exons anzugeben, wurden noch zwei weitere Kriterien mit einbezogen. Erstens sollten in den postulierten Exons durchgehende Leseraster existieren und zweitens wurde ein statistisches Prüfverfahren angewandt, das mögliche kodierende Sequenzen von Intronsequenzen unterscheiden sollte (112). Das Verfahren beruht auf dem Befund, daß viele Gene eine Präferenz für bestimmte Codons aufweisen und damit einzelne Basen an der jeweils dritten Stelle eines Triplets häufiger zu finden sind, als bei einer statistischen Verteilung der Basen zu erwarten wäre. Es wird dabei unterstellt, daß die Basenverteilung von Intronsequenzen einem solchen Muster nicht folgt. Beide Kriterien setzen natürlich voraus, daß überhaupt proteinkodierende Sequenzen vorhanden sind.

Bei der Anwendung dieses Prüfprogramms auf die Sequenz von pSt 608 lassen sich eine Reihe möglicher, kodierender

Bereiche angeben. Zusammen mit diesen Daten, den gefundenen Exon-Intron Übergängen und potentiellen Leserastern wurde die Lage hypothetischer Exons definiert. In der Tabelle 14 sind die Bereiche hoher statistischer Wahrscheinlichkeit von kodierenden Sequenzen und die Lage der möglichen Exons angegeben.

Tabelle 14 : Hypothetische Lage von Exons im Klon pSt 608
 Positionsangaben : s. Abb. 38
 ? : keine Aussage möglich
 / : alternative Möglichkeiten

| Exon | Potentiell kodierend | Lage der Übergänge |
|------|----------------------|-----------------------|
| 1 | 50 - 80 | - 108/151 |
| 2 | ? | 205 - 286 |
| 3 | ? | 328 - 391 |
| 4 | 380 - 410 | ? |
| 5 | 700 - 780 | 694 - 811/844 |
| 6 | 1090 - 1130 | 1095 - 1170 |
| 7 | 1570 - 1720 | 1545/1581 - 1662/1759 |
| 8 | 1770 - 1860 | 1759 - 1866 |
| 9 | 1950 - 2090 | 1967 - (2100) |
| 10 | 2140 - 2240 | ? |
| 11 | (2440 - 2650) | ? |

Ob diese Art von Analyse auch im Bereich des hohen GC-Gehalts anwendbar ist, bleibt fraglich, da eine vollständige theoretische Analyse dieses Verfahrens für alle denkbaren Sequenzen noch aussteht.

Vorhandensein kollagener Sequenzen

Bei der Hybridselektion mit Hilfe der p550-Sequenz, wurde gezeigt, daß zumindest eines der selektierten Proteine (160 kd) mit bakterieller Kollagenase spaltbar war (s.C.1.5.).

Die Sequenz des Klon pSt 608 wurde deshalb daraufhin abge-
sucht, ob bei der Translation kollagene Sequenzen entstehen
können. Bei der Übersetzung aller möglichen Leseraster fan-
den sich nur sehr kurze Folgen von Wiederholungen der für
Kollagene typischen Sequenz $(\text{Gly-X-Y})_n$. Auch die für die
Spaltung mit der Kollagenase nötige Minimalsequenz
(Pro-X-Gly-Pro) fand sich an einigen Stellen in verschiede-
nen Leserastern. Falls solche Sequenzen in Protein übersetzt
würden, könnten sie die Spaltung mit Kollagenase erklären.

7. Untersuchung der cDNA Klone p2200 und p1300

7.1 Konstruktion der cDNA Klone p2200 und p1300

Bei der Konstruktion und dem Screening knorpelspezifischer cDNA Klone (s.E.) wurden neben dem Klon p550 noch zwei weitere cDNA Klone isoliert: p1300 und p2200. Diese Klone enthalten ebenfalls Bam HI - Hind III Fragmente, eingebaut in pBR 322. Deren Expression ähnelt der von p550. Bisher war von diesen Klonen nur die Länge des klonierten Inserts und die Lage weniger Restriktionsstellen bekannt. Im Folgenden ist die Restriktionskartierung und Sequenzierung der Klone, die Kartierung der kodierenden genomischen Regionen und die Bestimmung der mRNA Länge durch Northernblotting beschrieben.

7.2 Restriktionskartierung von p2200 und Subklonierung von Fragmenten

Die Restriktionskarte des cDNA Klon p2200 mit den wichtigsten Spaltstellen ist in Abb. 40 gezeigt und wurde durch Einfach- und Mehrfachspaltungen des gesamten Klon oder durch partielle Spaltung des isolierten und endmarkierten Inserts bestimmt.

Um einen leichteren Zugang zu den verschiedenen Regionen des Klon zu haben, wurden verschiedene Fragmente in M13 Vektoren subkloniert. Mit den bereits früher beschriebenen Methoden (s.C.1.1.) wurde zuerst das gesamte Insert (ein Bam HI/Hind III Fragment von 1.9 kb Länge) sowohl in M13 mp8 als auch in M13 mp9 kloniert. Die rekombinanten Klone wurden als 2200 mp8 und 2200 mp9 bezeichnet. Neben der Sequenzierung

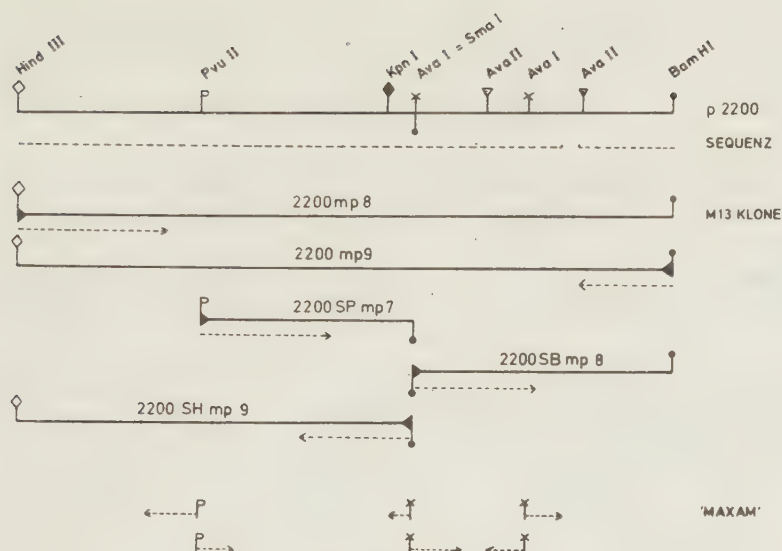


Abb. 40 : Restriktionskarte des cDNA-Klons p2200 und Lage der Subklone
 ----> : sequenzierte Bereiche
 Maxam : Sequenzierung nach der Methode von Maxam & Gilbert (markiertes 5'-Ende angegeben)

der Randsequenzen des cDNA-Klons, ist mit diesen Klonen wieder die Bestimmung der Orientierung der mRNA-Sequenz möglich, da nur einer der beiden komplementären Klone mit mRNA hybridisieren sollte.

Die nur einmal auftretenden Restriktionsstellen von Pvu II und Sma I dienen zur Konstruktion der drei Klone 2200SP mp7, 2200SH mp9 und 2200SB mp8 (Abb. 40). Die klonierten Fragmente des cDNA Klons, die zur Klonierung verwendeten Vektoren und die benutzten Klonierungsstellen sind in der Tabelle 15 zusammengefaßt.

Tabelle 15 : Klonierung von Fragmenten des cDNA Klons p2200
in M13 Vektoren

| Klon | Fragment | Länge | M13-Vektor |
|------------|-----------------|--------|----------------------|
| 2200mp8 | Bam HI/Hind III | 1.9 kb | mp8/Bam HI+Hind III |
| 2200mp9 | Bam HI/Hind III | 1.9 " | mp9/Bam HI+Hind III |
| 2200SP mp7 | Sma I/Pvu II | 0.6 " | mp7/Hinc II |
| 2200SH mp9 | Sma I/Hind III | 1.1 " | mp9/Hinc II+Hind III |
| 2200SB mp8 | Sma I/Bam HI | 0.75 " | mp8/Hinc II+Bam HI |

Da bei der Spaltung mit Sma I und Pvu II Fragmente mit 'Blunt Ends' entstehen, wurde in diesem Fall die Hinc II Stelle des Vektors zur Klonierung verwendet. Die geringere Effizienz einer Ligation solcher glatter Enden führte zu einer geringen Ausbeute an rekombinanten Klonen, besonders im Falle des Klons 2200SP mp7. In Abb. 40 ist die Lage und die übliche Sequenzierungsrichtung dieser Klone angegeben.

7.3 Sequenzierung des Klons p2200

Die Sequenzierung der M13 Subklone des cDNA Klons p2200 führte zur Bestimmung eines Großteils der Sequenz. In der Abb. 40 sind die Bereiche angegeben, deren Sequenz mit der Didesoxymethode bestimmt wurde. Alle Klone wurden mindestens zweimal in unabhängigen Ansätzen sequenziert. Zwei Bereiche der Sequenz konnten damit jedoch nicht erfaßt werden: die Umgebung der Pvu II-Stelle und die Region zwischen den beiden Ava II-Spaltstellen.

Die Sequenzierung dieser Bereiche erfolgte mit der Methode nach Maxam und Gilbert (64). Welche Fragmente isoliert, mit $\alpha^{32}\text{P}$ -ATP und Polynukleotidkinase markiert, nachgespalten und sequenziert wurden, ist der Tabelle 16 zu entnehmen.

Die angegebenen Fragmente, die nach der Nachspaltung nur noch an einem Ende markiert waren, wurden aus Polyacrylamidgelen isoliert und sequenziert. Die Auftrennung der Produkte

erfolgte in 80 cm langen, 0.4 mm dicken Sequenzgelen. Dabei konnten zwischen 100 und 200 Basen der verschiedenen Fragmente gelesen werden. Die Lage, Länge und Orientierung dieser Sequenzen ist in Abb. 40 angegeben.

Tabelle 16 : Sequenzierung von Fragmenten des Klons p2200
nach Maxam & Gilbert

| Markierte Fragmente
(Spaltung von p2200) | Nachspaltung | Sequenzierte Fragmente |
|---|--------------|------------------------|
| 2.85 kb Pvu II | Hind III | 0.53 kb |
| 3.05 kb Pvu II | Kpn I | 0.52 kb |
| 0.32 kb Ava I | Ava II | 0.20 kb/ 0.12 kb |
| 1.45 kb Ava I | Bam HI | 0.40 kb |
| 4.1 kb Ava I | Pvu II | 0.62 kb |

Diese zusätzlichen Sequenzen erlaubten die Vervollständigung der Sequenz im Bereich zwischen der Hind III und der Pvu II Stelle. Die Lücke zwischen den beiden Ava II Stellen konnte noch nicht vollständig sequenziert werden. Etwa 50 bp der Sequenz sind deshalb noch nicht erfaßt. Die Sequenzen bestätigten in den überlappenden Bereichen die vorher durch Sequenzierung der M13 Klone bestimmten Sequenzen. Alle Sequenzen sind in Abb. 41 zu einer Gesamtsequenz zusammengefaßt. Die Nummerierung der Positionen erfolgte unter der Annahme, daß die Länge des fehlenden Sequenzstücks etwa 50 bp beträgt.

Untersuchung der Gesamtsequenz von p2200

Sucht man die Sequenz des Klons p2200 auf das Vorhandensein eines durchgehenden Leserasters ab, so findet man unter den sechs Möglichkeiten keinen Raster, der ohne Unterbrechung in ein Protein übersetzt würde. Die nochmalige Überprüfung aller Sequenzgele auf eventuelle Lesefehler ergab kein anderes Ergebnis.

Abb. 41: Sequenz des cDNA Klones p2200

```

1   AAGCTTCAAC GTGACGGTGA GCAGCAGCAG CAGCCAGAAG CCCCCGCAGT
51  GTGAAGCAGA AAGGAGCGCT GCGGGTGCTG CCCAATCCAC CATCACCTTG
101 TCCAGGCAGC ACTTGGATGT GGCAGAGCCG CAGGGCTCCC CAGAGGAGCT
151 GGTGTACAGC CTGCTGAGCA GCCCCAGGC CCGACACCTG GCTAATGTGC
201 GCAACCCACA GGTGCCCATC AAGCGTTCA CACAGGCGGA CATCAATGCG
251 GGCCACGTGG TGTGGTTCAC TGCAGCCCGT GGTCTCAGG AGTCCCTGGC
301 ACTCAGCGTG TCTGACGGCC GGCAGGCAGG CGGGCGCATC TCACTGGAGG
351 TGGTGGTGCT ACCCACAGGC AGTGCAGAGC TGCTGCAGGT GCCGCTGGAG
401 CTGAAATGGT CCACGCTGAC CCACGGCAGC TGCTCCGCAG CTTGGGGAGG
451 CGATGCCCTT TACACCATCA CCACGGAGCC GAGGTTTGCC CAAGTGCAGG
501 TCAACAGAAA GCCAGCACAA GGCTTCTCGC AGGCCAGCTG GAACGCGGGG
551 AGGTGACCTT CACCTTCACC AAGCTCACCT CGTCCGAGGA CAGCTTCCAG
601 TTCCTACCA GGTCCCAGAC AATGAACAGC ACTGGGCTGG TGAACGTCAC
651 CGTCCCTGCC TTGGTGACAG CGAGGCCGGG CAGTGTGTGG CCCAGGGGAC
701 CACAGCCCTG CTGATACCA GACACTGGAT GCCAGTGGC TGCCAACCGC
751 ACCCAGACGA TCCCGGTCTT CCGCCTCCGA GACGCCCCCA CAGCAGCCGC
801 CTCATCAGGA GGCCCCGTGC GCCAGGACAG CCTGCCATCC CTATTGAATC
851 CTTCACTCAG AGGGAGCTGG AGCAAGGGCT GGTGGGGCTG GAGGTGCTGG
901 ATGACGGGGA CACAAACCAG CAGCACCCGC ACAGTGACAG CTTNACCTTT
951 GAGTTGGTGG CTGTGNNTGC ACNAGNANTG NTGTCTCTGG GGTACATCAC
1001 CGAGCCCTAC AATGCCTCTA AATCCTATGG TGTCACCCTG CTCCTGGCCC
1051 CACCAAGCAC ATCGTCCCTG GTACCACAGC CGCAGGGCAC AGCAGCGGAG
1101 AGCCCCAACG CCAGCGGAGCC AGCTCTGTGC CCCCACCCAC CAGCCCCAGC
1151 CCGGGGGAGG GCGGCACTTT CCTAGGCTTC ATCGAGGCCA ACATGTTCAG
1201 CATCATCATC CCACTCTGCC TCATCATCCT CCTGCTGGCC CTCATCCTGC
1251 CCCTGCTCTT CTACTGACAC AAGCGCAACA AGACGGGCAA GCACCACGTG
1301 CAGGGCTCGC CTGCCTCCAA GGCCAAGAAC GGGGCTGTAC CAGAGCAGGA
1351 GACCTTCCGC AAGACGGACC CCAACCAAGG CATCCCCCTA ACGACTGTCA
1401 ATGCCCTGGA TGGCAAAGGC ACAGCTCCAT CGCCGCAGGG TCCAGCGCTG
1451 CCGGCTGACC CCGAGCCGCT GCAGTACTGC CGGACATCCA ACCCCGCCCT
1501 GAAAAACAAC CAGTACTGGG TGTGAGGGCG CAGCGCGGGG GCACGAGTGG
1551 TGGGAGCCCC CCAGAGCGTC ACCTGCCT.. ..... (etwa 50 bp)
1601 .....TG GGGGCTCCAG CGTGGCCACC
1651 CAGGACCCCC CCGTCTCCAT AGTGCTTTGG GGCAGAGGGG CAAGTGGGTG
1701 GGGGGAGGCA AAACGCCATC AGGGGACGCT GCCATCGGGT ACGCAATGGG
1751 GGCACCTGAG GTTGGCCGAG GTGCTGGGGG TGCAGGCGGG TGAGAGGGAC
1801 GAGGACGGAG CACTTTGCGC AACGTCAGGC TCAGGATGGA GCTGCCCAAA
1851 TCCCCAGGGC TAGGCGGGCT GCAGTGACGG AGCAGAGGCT TGGAAAAACC
1901 CTGTGGCCAA CGCAAAGGGG ATCC

```

Die Sequenz des cDNA-Klons p2200 stellt damit keine durchgehende, proteinkodierende Sequenz dar. Eine Erklärungsmöglichkeit ist die Lokalisierung dieser Sequenz in den 5'- oder 3'-terminalen Bereichen der mRNA. Randsequenzen von p2200 könnten damit Beginn oder Ende der translatierten Sequenz enthalten. Als alternative Möglichkeit ist denkbar, daß die cDNA nicht an einer reifen mRNA gebildet wurde son-

dern an einer noch nicht vollständig prozessierten Vorlauferform. Die Sequenz von p2200 wurde damit teilweise oder vollständig Intronsequenzen entsprechen.

7.4 Genomorganisation des p2200 kodierenden Bereichs

Der Vergleich der Restriktionskarten des cDNA Klons und des entsprechenden kodierenden Bereichs des Genoms sollte Rückschlüsse auf die Exonstruktur zulassen und damit vielleicht indirekt eine Erklärung für den fehlenden Leseraster liefern.

Die Kartierung der Umgebung der p2200 kodierenden Region des Genoms erfolgte durch Spaltung genomischer DNA mit verschiedenen Restriktionsenzymen, Auftrennung der Fragmente durch Gelelektrophorese, Southern-Blotting und Hybridisierung mit dem markierten Insert von p2200. Alle genomischen Fragmente, die Teile der p2200 Sequenz enthalten, ließen sich dadurch nachweisen. Das Beispiel einer solchen Hybridisierung ist in Abb. 42 gezeigt. In der Tabelle 17 sind die Längen der hybridisierenden Fragmente verschiedener Spaltungen zusammengefaßt.

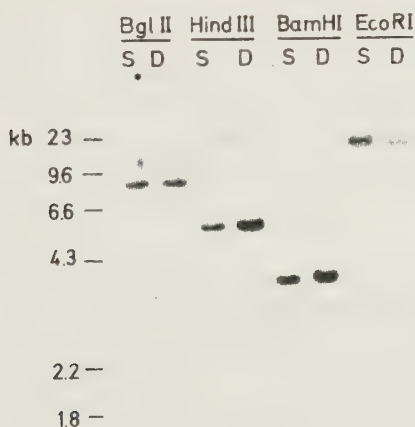


Abb. 42 :

Southerntransfer genomischer Spaltungen und Hybridisierung mit nick-translatiertem Insert von p2200 (Autoradiographie)

Je 10 µg DNA :

S = *Sterna*

D = Dedifferenzierete Chondrocyten

Gel : 0.8% Agarose

Tabelle 17 : Hybridisierung genomischer Spaltungen mit p2200

| Spaltung | Hybridisierende Fragmente (kb) |
|-------------------|--------------------------------|
| Eco RI | 20 |
| Bam HI | 3.9 |
| Hind III | 5.8 |
| Bgl II | 8.4 |
| Kpn I | 25 / 10.0 |
| Eco RI + Bam HI | 3.9 |
| Eco RI + Hind III | 5.8 |
| Bam HI + Hind III | 1.95 |
| Kpn I + Hind III | 4.6 / 1.05 |
| Kpn I + Bam HI | 3.0 / 0.84 / (3.8 partial) |

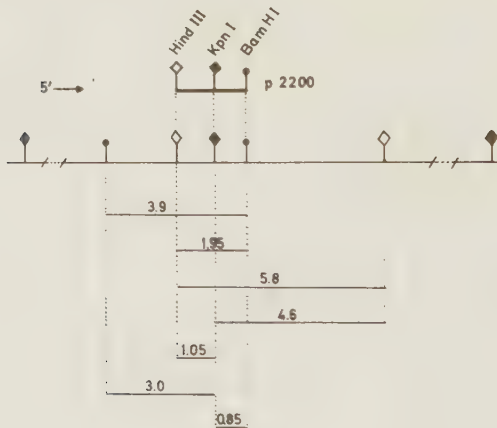


Abb. 43 : Restriktionskarte des p2200-kodierenden genomischen Bereichs (Fragmentlängen in kb)

Aus den Fragmenten der Tabelle 17 ließ sich die Restriktionskarte der p2200 kodierenden genomischen Region ableiten, die in Abb. 43 gezeigt ist. Die Lage der einzelnen hy-

bridisierenden Fragmente ist dort angegeben. Lediglich eine genaue Lokalisierung der weiter entfernten Eco RI- und Kpn I-Spaltstellen war nicht möglich, da die Daten der Spaltungen dafür nicht ausreichten.

Der Vergleich der genomischen Restriktionskarte mit der des cDNA-Klons zeigte keinen Unterschied in der Länge des Bam HI/Hind III-Fragments und in der Lage der Kpn I-Stelle. Der cDNA Klon entspricht damit genau der kodierenden Sequenz im Genom, mit der Einschränkung, daß sehr kleine Längendifferenzen mit dieser Methode nicht unterscheidbar sind.

Damit sind zwei Deutungen der Sequenz des Klons p2200 denkbar: entweder repräsentiert der Klon die Sequenz eines Exons, das keine oder nur teilweise kodierende Sequenzen enthält, oder die cDNA wurde an einer noch nicht prozessierten, stabilen Vorläuferform der mRNA synthetisiert.

7.5 Orientierung und Expression des p2200 Gens

Bestimmung der Orientierung der p2200 Sequenz

Aus den Sequenzdaten alleine war es nicht möglich, den kodierenden Strang zu bestimmen. Aus diesem Grund wurden die beiden komplementären Klone 2200 mp8 und 2200 mp9 zur Bestimmung der Orientierung der Sequenz in der mRNA verwendet. Beide Klone wurden radioaktiv markiert und mit RNA hybridisiert. Bei der Dothybridisierung von RNA embryonaler Sterna (17 Tage) und Kalvarien fand sich nur Hybridisierung mit 2200 mp8 (ohne Abb.). Ähnlich der Ableitung der Orientierung von p550 (s. Abb. 12C), ergab sich im Falle des cDNA-Klons p2200 folgende Anordnung auf der mRNA : 5'-Hind III-Bam HI-3'. Die Gesamtsequenz von p2200 in Abb. 41 ist bereits in dieser Orientierung angegeben.

Ausschließlich die RNA aus Sterna hybridisierte mit 2200 mp8, nicht jedoch die RNA aus Kalvarien. Um dieses Ergebnis zu bestätigen, wurden die gleichen RNA Präparationen in

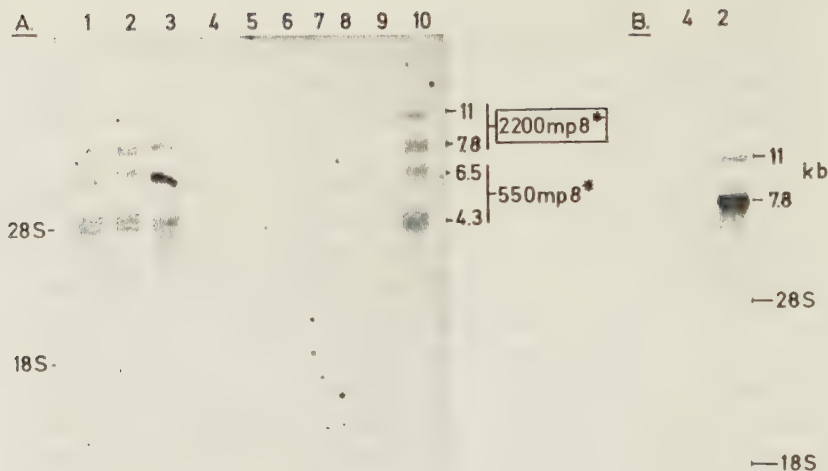


Abb. 44 : Northernhybridisierung mit verschiedenen gewebe-spezifischen RNAs (Autoradiographien)

1 : Sterna/ 12 Tage Embryonen

2 : Sterna/ 19 "

3 : Kalvarien/ 12 "

4 : Kalvarien/ 19 "

5 : Embryofibroblasten

6 : Leber/ 17 "

7 : Hirn/ 17 "

8 : Epiphysen/ 19 "

9 : Röhrenknochen/ 19 "

10: Sehne/ 16 "

A. Gel : 1% Agarose / Filter identisch mit Abb. 13

1. Hybridisierung mit markiertem 550mp8

2. Hybridisierung mit markiertem 2200mp8

B. Gel : 0.8% Agarose

Hybridisierung nur mit markiertem 2200mp8

Agarosegelen aufgetrennt, durch Northernblotting transfe-riert und mit den beiden markierten Klonen hybridisiert. Abb. 44B zeigt nur die Hybridisierung mit 2200 mp8. Die hybridisierende mRNA hatte eine Länge von etwa 7.8 kb und fand sich ausschließlich in der Sterna-RNA. Zwei weitere, in geringer Menge auftretende Banden von etwa 9 und 11 kb Länge ließen sich zusätzlich erkennen.

Der cDNA-Klon p2200 repräsentiert damit eine sehr große, anscheinend spezifisch im Sternum exprimierte mRNA. Ob diese mRNA auch in anderen Geweben - ähnlich der p550 mRNA (s.C.1.4.) - zu finden ist, mußte noch getestet werden.

Gewebespezifität der Expression des p2200 Gens

Durch Northernblotting der RNA verschiedener Gewebe und Hybridisierung mit markierter DNA von 2200 mp8 ließ sich ein genaueres Bild von der Expression des p2200 Gens erhalten. In Abb. 24A ist die Autoradiographie einer solchen Hybridisierung gezeigt. Das verwendete Filter war bereits vorher mit 550 mp8 hybridisiert worden (s.Abb. 13). Die gebundene 550 mp8-Probe wurde abgewaschen, wobei sich ein Teil nicht mehr entfernen ließ, da durch Trocknen der Filter die Probe bereits fest fixiert war. Dies bot jedoch die Möglichkeit, die Verteilungsmuster der mRNAs, spezifisch für p550 und p2200, miteinander zu vergleichen.

Die mRNA Species von 7.8 kb Länge, die bereits in Abb. 44B gefunden wurde, ließ sich in der RNA sowohl von Sterna, als auch von den Kalvarien 12 Tage alter Embryonen nachweisen. Ebenso tritt sie in großer Menge in der RNA aus Sehnen auf. Mit der RNA aus Kalvarien 19 Tage alter Embryonen findet sich keine Hybridisierung mehr und bestätigte das Ergebnis der Dothybridisierung. Eine zusätzliche mRNA Species von etwa 11 kb findet sich in besonders großer Menge in der RNA von Sehne und nur in geringer Menge in der RNA von Kalvarien/ 12 Tage und Sterna/ 12 Tage.

Das Verteilungsmuster der p2200-spezifischen mRNA ähnelt sehr stark dem Muster der p550-spezifischen mRNA, mit dem einzigen Unterschied, daß die p2200 spezifische mRNA nicht mehr in den Kalvarien 19 Tage alter Embryonen zu finden ist. Ob ein Zusammenhang zwischen den beiden cDNA-Klonen p550 und p2200 besteht, läßt sich mit diesen Ergebnissen noch nicht entscheiden. Denkbar ist, daß beide cDNA Klone zwei unterschiedliche, unabhängige Gene repräsentieren, deren Expression koordiniert reguliert wird.

7.6 Restriktionskartierung und Sequenzierung des cDNA Klons p1300

Der cDNA-Klon p1300 enthält ein Insert von 1.25 kb, ebenfalls eingebaut in die Bam HI- und Hind III-Stelle von pBR 322. Die Karte der wesentlichsten Restriktionsenzyme dieses Klons zeigt Abb. 45.

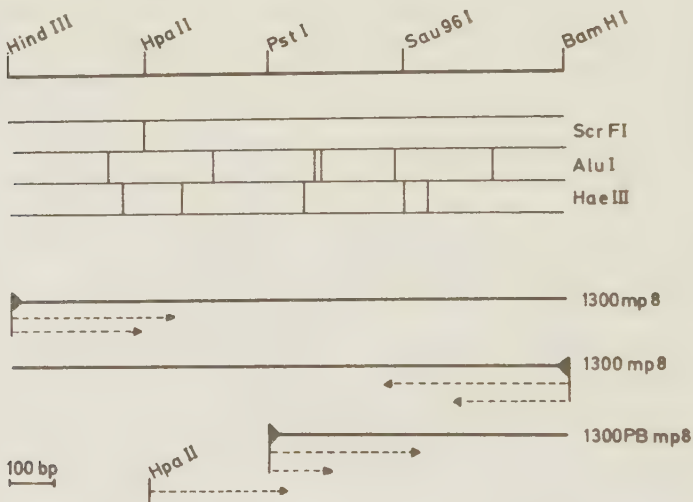


Abb. 45 : Restriktionskarte des cDNA-Klons p1300 und Lage der Subklone
 ---> : Sequenzierte Bereiche

Entsprechend den Klonen p550 (s.C.1.1.) und p2200 (s.C.7.2.) wurde das gesamte klonierte Fragment in die Vektoren M13 mp8 und mp9 eingebaut und die Klone mit 1300 mp8 und 1300 mp9 bezeichnet. Zusätzlich wurde das 0.65 kb lange Pst I/Bam HI-Fragment in M13 mp8 kloniert (1300PB mp8). Mit Hilfe dieser drei Klone konnte ein Großteil der Sequenz des cDNA-Klones bestimmt werden. Lage und Orientierung der Subklone und der gelesenen Sequenzbereiche sind in Abb. 45 an-

gegeben.

Die Sequenz in der Umgebung der Pst I Spaltstelle wurde zusätzlich nach der Methode von Maxam & Gilbert bestimmt. Das isolierte 0.95 kb Hpa II Fragment einer Spaltung des gesamten Plasmids wurde an den 5'-Enden markiert und mit Bam HI nachgespalten. Das entstehende 0.95 kb Hpa II-Bam HI Fragment wurde isoliert und sequenziert. Die Orientierung und Länge dieser Sequenz ist ebenfalls in Abb. 45 angegeben.

Gesamtsequenz von p1300 und Leseraster

Aus allen Einzelsequenzen ließ sich die Gesamtsequenz des Klons p1300 zusammenfügen. Die 1226 bp lange Sequenz ist in Abb. 46 gezeigt. Unsicherheiten in der gelesenen Sequenz, nicht jedoch im Leseraster, fanden sich nur an einigen Positionen im Bereich 400-550.

Sucht man diese Sequenz nach möglichen proteinkodierenden Leserastern ab, findet sich nur ein durchgehend, translatierbarer Raster, dessen Übersetzung ebenfalls in der Abb. 46 angegeben ist. Eine Proteinsequenz von 408 Aminosäuren ließ sich ableiten, die ein Teilfragment des p1300-Genprodukts darstellt. Die Aminosäurezusammensetzung dieses Teilstücks des p1300-Genprodukts ist in der Tabelle 18 angegeben.

Tabelle 18 : Aminosäurezusammensetzung eines Teilfragments
des p1300 Genprodukts
Angaben : /1000 Aminosäuren

| | | | | | |
|-----|-----|-----|----|-----|----|
| Asn | 34 | Gly | 22 | Tyr | 25 |
| Asp | 17 | Ala | 54 | Phe | 10 |
| Thr | 108 | Cys | 22 | His | 59 |
| Ser | 100 | Val | 64 | Lys | 47 |
| Gln | 69 | Met | 29 | Arg | 39 |
| Glu | 42 | Ile | 64 | Trp | 2 |
| Pro | 56 | Leu | 59 | X | 59 |

Abb. 46 : Sequenz des cDNA Klonen p1300 und Übersetzung des
einzigen durchgehenden Leserasters
(Aminosäure unter der 1. Base eines Codons)

| | 10 | 20 | 30 | 40 | 50 |
|------|-------------|------------|------------|-------------|------------|
| 1 | AAGCTTTTCGG | AAAGCAAGAG | TAACTAATCA | CAAGAAATGGA | ATTAATCTCC |
| | S F R | K A R V | T N H | K N G | I N L Q |
| 51 | AGACAGTAAC | AACTAGGCCA | GATAAACACC | TTCTTACAGA | TACTGGCTAT |
| | T V T | T E P | D K H L | P T D | T G Y |
| 101 | TACACTAATG | AGAAGTCTGC | CCAGAAGATA | ATGTCTTCTC | AGATATCTAA |
| | Y T N E | K S A | Q K I | M S S Q | I S N |
| 151 | CACAAATGTG | TCCTATGTAC | AAACAATCAC | TAGACCACTG | ACCACTATCT |
| | T N V | S Y V Q | T I T | R P L | T T I S |
| 201 | CCAGTGTAGG | CACAAGCACA | GTAGCTGGTA | TTAGTGCTAC | TCTCCATCAA |
| | S V G | T S T | V A G I | S A T | L H Q |
| 251 | GGCCAAATGC | AGAAGTCTCA | AAAGCGGAGA | GTATCTACCC | TTAAATTAAC |
| | G Q M Q | K S Q | K R R | V S T L | K L T |
| 301 | CCGGACACAA | GAATCGGCAA | GTAGTGTTCA | CTGTAACATA | CAGCAACAGC |
| | R T Q | E S A S | S V H | C N I | Q Q Q x |
| 351 | NCAAGCCTGT | TGAAGTCCTT | ATATTTTNNC | AAAGGCCTGT | CNNNCCTGCC |
| | K P V | E V P | I F x Q | R P V | x P A |
| 401 | TGTAACCNCA | TGGCTGAGAT | GCACACAGCA | TCNNNTTNA | TTNNGATAGC |
| | C N x M | A E M | H T A | S x x I | x I A |
| 451 | TNNNATNNCT | GACAATAATG | CAGAGCCTGA | AACNCAAGAG | CATNNACCTA |
| | x x x | D N N A | E P E | T Q E | H x P S |
| 501 | GTATTNNNN | NGACACTTNN | NAGATGTCTT | NNNAGGATGT | NNNACAAAAG |
| | I x x | D T x | x M S x | x D V | x Q K |
| 551 | TCTACAGTTN | NZNNNNNNNN | NNNGGCTGCA | GAATGCCATG | TGGCTCAGAT |
| | S T V x | x x x x | x A A | E C H V | A Q I |
| 601 | CACAGTGAAT | GGGAACCTCG | GCAGTGCTGC | CAGCCCATG | AGTCACCTTC |
| | T V N | G N S G | T A A | S P V | S H F Q |
| 651 | AAAGGCCCTT | TTCCCTTCT | TCAGCTTACT | CTCCACCAGC | TTCTCTCAAT |
| | R P F | S P S | S A Y S | P P A | S L N |
| 701 | TCCAACCTTG | TTATCATGCA | GCATGGCAGG | ATGATGGAGT | CCACAGAGAC |
| | S N L V | I M Q | H G R | M M E S | T E T |
| 751 | ATACTCGCAA | CATGTTTACA | CCGTTGGCAG | CAGCACCACC | ACCAGTACAA |
| | Y S Q | H V Q T | V G S | S T T | T S T I |
| 801 | TACCAATCTG | TAAATCCTTA | GAGGAAGAAA | AAAAAATTAC | AGTCATTAAA |
| | P I C | K S L | E E E K | K I T | V I K |
| 851 | GCTCCACATT | ATGCAGGGAT | TGGCCCAGTG | GATGAATCTG | GAATCCCTAC |
| | A P H Y | A G I | G P V | D E S G | I P T |
| 901 | AGAATTAGAA | CGACAGTTGA | CAGGCCAAAG | GATTGGTACA | AGACAATGTT |
| | E L E | R Q L T | G Q T | I G T | R Q C S |
| 951 | CAAGCAAATC | CATATGGTAC | ATAAGCCTGA | TGACGACACA | CACATGTATA |
| | S K S | I W Y | I S L M | T T Q | T C I |
| 1001 | ATACTGCATA | TTCATACAGT | ACTGGTGGCT | ACAGTCCATC | CTTCTCTGCT |
| | I L H I | H T V | L V A | T V H P | S L L |
| 1051 | CAGGCACACC | CAGCTGCAAA | GACACAGACA | TACAGACCAT | TGTCTAAAAG |
| | R H T | Q L Q R | H R H | T D H | C L K V |
| 1101 | TGCTTCTGAC | AGCGGCTCCG | ATGCCTTTAA | GGTCTCATCT | CCTTTACCAC |
| | L L T | A A P | M P L R | S H L | L Y H |
| 1151 | CCCCACATGT | GCCACCCCCA | GTACCACCAC | NTCGTCTGCG | GGAAAGGTCT |
| | P H M C | H P P Q | Y H H | x V C G | K G L |
| 1201 | ATACCAGAAA | AAAATGACTG | GGATCC | 1226 | |
| | Y Q K | K M T G | I | | |

Die DNA Sequenz und die abgeleitete Proteinsequenz wurden wieder dazu verwendet, in den verfügbaren Datenbanken nach Homologien mit bereits bekannten DNA- oder Proteinsequenzen zu suchen. Es ließ sich jedoch keine bereits bekannte Sequenz finden, die der p1300 Sequenz entsprach. Das Genprodukt des p1300-Gens repräsentiert damit ein bisher nicht beschriebenes Produkt. Auch aus dem Vergleich der Daten der Aminosäurezusammensetzung mit denen bekannter Proteine (besonders der Knorpelmatrix) brachte keinen Aufschluß über die Natur dieses Proteins.

Orientierung der p1300 Sequenz

Die Orientierung der Sequenz von p1300 wurde wieder durch die Hybridisierung von RNA mit den beiden komplementären Klonen 1300 mp8 und 1300 mp9 ermittelt. Nur 1300 mp8 hybridisierte in einem Dohybridisierungsexperiment mit Sterna-RNA. Damit leitet sich folgende Anordnung auf der mRNA ab: 5'-Hind III-Bam HI-3'. Die Sequenz in Abb. 46 wurde bereits in dieser Orientierung angegeben. Der gefundene Leseraster ist damit zwar nicht bewiesen, aber sehr wahrscheinlich geworden.

7.7 Kartierung der p1300 kodierenden genomischen Region

Die Restriktionskarte der p1300 kodierenden genomischen Region wurde auf ähnliche Weise bestimmt, wie es bereits für den Klon p2200 beschrieben wurde (s.C.7.4.). Durch die Hybridisierung genomischer Spaltungen mit dem nick-translatierten Insert von p1300 wurden alle Fragmente erfaßt, die Teile von p1300 enthalten. Die Längen dieser hybridisierenden Fragmente sind in der Tabelle 19 angegeben und die daraus abgeleitete Restriktionskarte ist in Abb. 47 gezeigt.

Die Lage der Sequenz des cDNA-Klons p1300 läßt sich auf ein 1.25 kb Bam HI/Hind III-Fragment einengen, das genau der Länge des Inserts des cDNA Klons entspricht. Aus dieser

Parallelität von Gen und cDNA bzw. mRNA ergibt sich, daß in diesem Bereich keine Introns enthalten sind. Im Unterschied zu p2200 enthält die Sequenz jedoch einen durchlaufenden Leseraster, was die Vermutung nahelegt, daß der cDNA Klon p1300 einem Teil eines einzelnen proteinkodierenden Exons entspricht.

Tabelle 19 : Hybridisierung genomischer Spaltungen mit p1300

| | | |
|-----------------|------|--------------------|
| Eco RI | 3.8 | kb |
| Bam HI | 6.5 | " |
| Hind III | 1.3 | " (3.4 kb partial) |
| Bam HI+Hind III | 1.25 | " |
| Bam HI+Eco RI | 2.05 | " (3.5 kb partial) |
| Eco RI+Hind III | 1.3 | " |

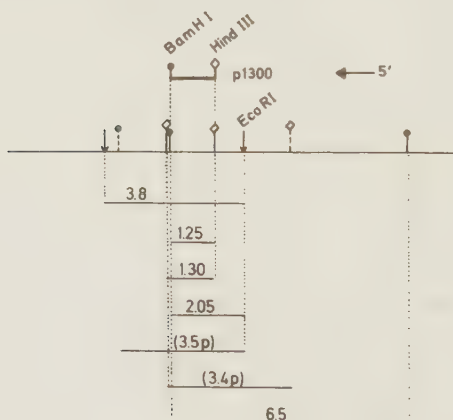


Abb. 47 : Restriktionskarte des p1300 kodierenden genomischen Bereichs

7.8 Expression der p1300 spezifischen mRNA

Durch Northernblotting von mRNA, isoliert aus *Sterna* und Kalvarien 17 Tage alter Embryonen, und Hybridisierung mit dem markierten Klon 1300 mp8 konnte gezeigt werden, daß diese cDNA - ähnlich p550 (s.C.1.4.) - keine absolute Spezifität für Knorpelgewebe aufweist. Wie Abb. 48 zeigt, finden sich, wenn auch in geringer Menge, hybridisierende mRNAs auch in der RNA von Kalvarien. Die Sequenz von p1300 findet sich dabei nur in einer mRNA Species von 4.3 kb Länge wieder. Zusätzliche Banden lassen sich in Abb. 48 nicht erkennen, anders als bei p550 und p2200, die beide eine Hybridisierung mit zwei verschieden langen mRNA-Species zeigten. Vorläufige Ergebnisse der Hybridisierung von p1300 mit der RNA verschiedener Gewebe deuten auf ein Verteilungsmuster hin, das große Ähnlichkeit mit dem des cDNA-Klons p550 aufweist. Dies muß jedoch noch durch weitere Versuche bestätigt werden.

| | | |
|---------------|----------|----|
| <u>K</u> | <u>S</u> | |
| 20 10 20 10 5 | | µg |

Abb. 48 :
Northernhybridisierung verschiedener Mengen RNA mit markiertem 1300 mp8*

S : *Sterna*/ 17 Tage
K : Kalvarien/ 17 Tage
Gel : 1% Agarose

28S —  — 4.3kb

18S —

Aufgrund dieser Ähnlichkeit wurden die genomischen Klone des p550-Gens (λ St 550-3 und λ St 550-5) nach Sequenzen abgesucht, die der des Klon p1300 entsprechen. Die Klone wurden dazu mit Restriktionsenzymen gespalten, auf Filter transferiert und mit dem nicktranslatierten Insert von p1300 hybridisiert. Es fand sich kein Fragment, das Teile dieser cDNA-Sequenz enthielt. Das p550- und das p1300-Gen repräsentieren damit wahrscheinlich unabhängige Gene, wobei unterstellt wird, daß auch keine Ähnlichkeit mit bisher nicht isolierten Bereichen zu finden ist. Möglicherweise gehören alle untersuchten cDNA-Klone (p550, p1300 und p2200) zu einer Familie koordiniert exprimierter Gene (s.G.6.).

D. DISKUSSION

Während die verschiedenen Kollagene, die strukturgebenden Proteine der meisten Bindegewebe, sowohl auf der Ebene der Proteine, als auch auf der Ebene der Gene bereits in vielen Einzelheiten beschrieben wurden, ist über andere Komponenten (Proteoglykane, Glykoproteine) dieser Gewebe nur sehr wenig bekannt. Einer proteinchemischen Charakterisierung standen teils der komplexe Aufbau, teils die geringen Mengen der einzelnen Proteine im Weg.

Grundgedanke dieser Arbeit war es, durch Konstruktion und Charakterisierung von c-DNA- und genomischen Klonen einen anderen experimentellen Zugang zu diesen 'amorphen' Anteilen im Knorpelgewebe zu erhalten. Damit sollte es möglich werden, die Struktur dieser Proteine zu beschreiben, sowie ihre funktionelle Bedeutung und ihre entwicklungsabhängige Expression verfolgen zu können. Da die gängigen Wege der Identifizierung eines Klons (Sequenzvergleich von DNA und Protein; immunologische Identifizierung) auf Grund der fehlenden Informationen unmöglich war, beinhaltet die weitere Untersuchung dieser Klone die Suche nach kodierenden Sequenzen, die zur Herstellung spezifischer Antikörper verwendet werden können.

1. Identifizierung des p550-Genprodukts

Der erste Teil der vorliegenden Arbeit befaßte sich mit dem c-DNA Klon p550. Seine Sequenz wurde vollständig ermittelt, die Größe der kodierten Proteine durch Hybrid-Selektions - Translations Experimente bestimmt und die gewebe- und entwicklungsabhängige Expression der spezifischen mRNAs verfolgt. Die grundlegenden Beobachtungen lassen sich in den folgenden Punkten zusammenfassen :

- (1) Es gibt einen durchgehenden Leseraster, der in seiner Orientierung der Leserichtung der mRNA entspricht.
- (2) Aus der cDNA Sequenz läßt sich die Proteinsequenz eines Teilfragments des p550 Genprodukts ableiten.
- (3) Isolierte mRNA läßt sich in drei Proteine unterschiedlicher Länge übersetzen : 160 kd (12 Tage/Sterna) und 140, 130 kd (16 Tage/Sehne und 19 Tage/Sterna).
- (4) Das 160 kd Protein ist durch bakterielle Kollagenase spaltbar.
- (5) Zwei mRNA Species, 4.3 und 6.5 kb lang, enthalten Teile der cDNA-Sequenz.
- (6) Die Expression des p550 Gens findet sich nicht nur im hyalinen Knorpel der Sterna, sondern auch in Sehnen und Kalvarien.
- (7) Während der Embryonalentwicklung (12-19 Tage) nimmt der Gehalt an p550 spezifischer mRNA in den Kalvarien ab, nicht jedoch in Sehnen und Sterna.
- (8) Bei der Dedifferenzierung von Chondrocyten wird die Expression des Gens ebenfalls abgeschaltet (44).

Aus diesen Befunden ergeben sich im wesentlichen drei Fragestellungen :

- Für welches Produkt kodiert das p550 Gen?
- Welche Funktion hat das Genprodukt in unterschiedlichen Geweben?
- Wie läßt sich das Muster der mRNA Expression erklären?

Die Ergebnisse bei der Untersuchung des cDNA Klons p550 geben nur teilweise Antworten auf die ersten beiden Fragen.

Eine detaillierte Beschreibung des Expressionsmusters des p550 Gens ist dagegen erst in Zusammenhang mit dessen Genstruktur möglich und wird deshalb später diskutiert (s.D.5.).

Sequenzvergleich

Die sicherste Identifizierung eines Proteins ist über den Vergleich von Proteinsequenzen möglich. Die abgeleitete Proteinsequenz des cDNA-Klons p550 wurde deshalb mit bekannten Proteinsequenzen einer Datenbank verglichen. Dabei fanden sich keine Proteine, die homologe Sequenzen enthielten. Ebenso wenig ließen sich auch beim Vergleich der cDNA-Sequenz mit zugänglichen DNA-Sequenzen verschiedener Datenbanken keine Homologien feststellen. Mit dieser Art von Vergleich war es also nicht möglich, das p550-Genprodukt einem bereits beschriebenen Protein zuzuordnen. Ein solches Ergebnis war jedoch nicht überraschend, da mit Ausnahme der Kollagene und des Fibronektins keine Sequenzen von Matrixkomponenten in den Datenbanken vorhanden sind.

Die Aminosäurezusammensetzung des sequenzierten Teilfragments des p550-Proteins (hoher Gehalt an Glu, Lys und Ile; wenig Gly) ist wenig hilfreich für eine eindeutige Identifizierung. Eine vergleichbare Zusammensetzung (hoher Gehalt an Dikarboxyl- und aliphatischen Aminosäuren) findet sich häufig in 'Strukturellen Glykoproteinen' einer Reihe von Geweben (113), deren einzelne Komponenten jedoch nicht im Detail beschrieben sind. Die Zuordnung zu dieser Gruppe ist auch deshalb recht unsicher, da die Zusammensetzung eines Teilfragments nicht notwendigerweise repräsentativ für das gesamte Protein sein muß.

Obwohl die abgeleitete Proteinsequenz keinen direkten Hinweis auf die Identität des Genprodukts gab, lieferte sie doch die Grundlage für einen weiterführenden Versuchsansatz zur Identifizierung des Genprodukts. In einigen Fällen konnte bereits gezeigt werden, daß kurze synthetische Oligopeptide (>7 Aminosäuren) ausreichen, spezifische Antiseren

gegen das gesamte Protein herzustellen (114). Aus diesem Grund wurden Sequenzen der p550 Teilsequenz ausgewählt, die mit großer Wahrscheinlichkeit auf der Außenseite des Proteins lokalisiert sein sollten, potentiell also als antigene Determinanten dienen könnten (115,116). Zwei dieser Sequenzen wurden ausgewählt (Pos. 124-132; Pos. 153-163; Abb. 11) und die entsprechenden Oligopeptide (Länge 9 und 11 Aminosäuren) synthetisiert. Mit Hilfe dieser Peptide soll versucht werden, spezifische Antiseren herzustellen, die zur Identifizierung des Gesamtproteins führen müßten (Immunpräzipitation, Western-Blotting, Immunfluoreszenz).

Expressionsmuster des p550 Gens

Weitere Informationen über charakteristische Merkmale des p550-Genprodukts ließen sich aus dem Muster seiner gewebe-spezifischen Expression ableiten (s.C.1.4.). Zum cDNA-Klon p550 komplementäre mRNA fand sich in Sterna, Sehnen und Kalvarien. Dies war überraschend, da in früheren Versuchen keine Hybridisierung mit Kalvarien-mRNA gefunden werden konnte (44). Die wahrscheinliche Ursache für diesen Unterschied liegt in den dort verwendeten Hybridisierungsbedingungen (54°C , 60% Formamid); sie täuschten eine spezifische Hybridisierung mit mRNA aus Sterna vor, da die geringeren Mengen in der RNA aus Kalvarien (19 Tage) unter diesen Bedingungen nicht mehr nachweisbar waren.

Interessant am gefundenen Expressionsmuster ist die Tatsache, daß die mRNA nicht einheitlich in einem einzigen Gewebetyp (z.B. Knorpel) zu finden ist, sondern in unterschiedlichen, wenig verwandten Geweben. Während im hyalinen Knorpel der Sterna spezifische mRNAs auftreten, konnten sie im epiphysealen Knorpel (von Fibula und Tibia) nicht oder nur in sehr geringen Mengen nachgewiesen werden. In den mineralisierten Anteilen dieser Röhrenknochen ist die mRNA ebenfalls nicht nachweisbar, wohl aber in den Kalvarien, zumindest vor oder zu Beginn der Kalzifizierung. Ausgehend von histologischen und embryologischen Beobachtungen war bisher

angenommen worden, daß zwischen Sterna und Kalvarien keine Ähnlichkeit besteht, da Kalvarien keine knorpelige Phase in ihrer Entwicklung durchlaufen (endesmale Ossifikation). Die Osteoblasten entstehen direkt aus den noch undifferenzierten Osteoprogenitorzellen (117). Diese Zellen könnten deshalb für die Synthese des p550-Produkts verantwortlich sein, da in späteren Stadien nur mehr eine geringe Anzahl dieser Zellen vorhanden ist, parallel dazu aber auch die Menge an p550-spezifischer mRNA abnimmt. Im Gegensatz zu den Kalvarien erfolgt im Falle der Röhrenknochen die Knochenbildung über eine knorpelige Zwischenstufe (enchondrale Ossifikation), die sich anscheinend wesentlich vom hyalinen Knorpel des Sternums unterscheidet, obwohl in beiden Fällen der gleiche Kollagentyp (Typ II) gefunden wird.

In der Phase der Kalvarienentwicklung vor oder während der Kalzifizierung werden dem Sternum vergleichbare funktionelle Ansprüche an das Gewebe gestellt: beide Gewebe benötigen eine gewisse Grundelastizität. Diese Eigenschaft geht im Verlauf der weiteren Entwicklung bei den Kalvarien verloren. Damit läßt sich die Hypothese aufstellen, daß das p550-Genprodukt eine Komponente elastischer Strukturen dieser Gewebe darstellt.

Solche elastischen Elemente wurden bisher im hyalinen Knorpel nicht beschrieben, ebenso wenig in Kalvarien. Elastin und Elastin-assoziierte Mikrofibrillen, die Hauptkomponenten der meisten elastischen Strukturen, konnten bisher nur in den elastischen Knorpeln (Epiglottis, Außenohr), nicht jedoch im hyalinen Knorpel (118) und in Kalvarien nachgewiesen werden. Denkbar ist, daß das p550-Genprodukt an vergleichbaren, strukturell unabhängigen Elementen beteiligt ist. Für eine solche Beteiligung spricht auch die Tatsache, daß große Mengen der spezifischen mRNA auch in Sehnen gefunden werden konnte, deren wesentliche Eigenschaft auf einer begrenzten Flexibilität und Elastizität beruht. Zumindest für die Sehne des Flügels konnte gezeigt werden, daß auch hier elastische Mikrofibrillen vorhanden sind (119).

Für das p550-Gen kann somit der Einfluß lokaler, biome-

chanischer Faktoren, die zu einer funktions-, nicht aber gewebeabhängigen Expression führen, als entscheidend angesehen werden. Um eine solche These zu überprüfen, müßte die Expression in einer Reihe weiterer Gewebe getestet werden (Aorta, Muskel etc.). Bei bestimmten Knorpeltypen konnte ein solcher Einfluß mechanischer Belastungen auf die Differenzierung des Gewebes bereits gezeigt werden (120).

Größe der kodierten Proteine

Ein direkter Zugang zum translatierten Protein selbst war durch die Hybridselektion spezifischer mRNA und deren Translation in vitro möglich (s. C.1.5.). Dabei ließ sich die mRNA in drei Proteine von unterschiedlicher Größe übersetzen (160, 140 und 130 kd). Dieses Ergebnis ist in mehreren Punkten überraschend :

- Zwei mRNA-Species (4,3 und 6,5 kb) führen zur Synthese von drei Proteinen.
- Es werden unterschiedliche Proteine von der mRNA aus Sterna/ 12 Tage und Sterna/ 19 Tage translatiert.
- Es besteht aber kein Unterschied in den Längen der mRNAs verschiedener Gewebe oder Entwicklungsstadien.

Da zwei mRNA Species zur Synthese von drei unterschiedlichen Proteinen führen, muß man annehmen, daß drei individuelle mRNAs in zwei Längenklassen vorliegen. Nicht so leicht erklären läßt sich die ausschließliche Synthese des 160 kd Produkts bei mRNA aus Sterna/ 12 Tage. Die mRNAs der 130 und 140 kd Proteine müssen in einer nicht translatierbaren Form vorliegen, ohne sich jedoch in ihrer Länge wesentlich zu unterscheiden. Denkbar wären Unterschiede in finalen Schritten des Splicings, die zu einer Aktivierung bzw. Inaktivierung einzelner mRNA Species führten. Demnach würde eine Steuerung der Expression einzelner Proteine auf

der Ebene des Splicings vorliegen.

Eine detailliertere Erklärung wird erst möglich sein, wenn die Vorgänge bei der Transkription und beim Splicing in allen Schritten geklärt sein werden. Die Versuche zur Struktur des p550-Gens gaben zwar einige Hinweise auf den Zusammenhang Gen - mRNA, eine eindeutige Zuordnung der kodierten Proteine zu einzelnen mRNAs war damit aber noch nicht möglich (s.G.5.).

Da nun sowohl die Längen der kodierten Proteine, das Expressionsmuster in verschiedenen Geweben und eine partielle Aminosäurezusammensetzung bekannt war, wurden diese Daten mit verschiedenen, in jüngster Zeit beschriebenen Matrixkomponenten genauer verglichen. In Tabelle 19 sind einige in Betracht kommende Komponenten aufgeführt.

Tabelle 19 : Vergleich des p550 Genprodukts mit minoren Matrixkomponenten

| Protein | p550 | PG-Lt | G1/G2 | CL | MFP I |
|----------------------------|-------------------------------|----------------------|----------------|--------------------------|--------|
| Literatur | | (25) | (121) | (122) | (123) |
| Species | Huhn | Huhn | Huhn | Rind | Rind |
| Gewebe | Sternum
Sehne
Kalvarien | Sternum
Epiphysen | Aorta
Sehne | Aorta
Sehne
Bänder | Bänder |
| Molekular-
gewicht (kd) | 160
140
130 | 120
100 | 140
130 | 140 | 150 |
| Kollagenase-
spaltung | 160 +
140/130 ? | + | - | + | + |
| | | | | (3 Fragmente) | |

Von allen diesen Proteinen ist bisher nur die Aminosäurezusammensetzung bekannt, nicht jedoch deren Proteinsequenzen. Gemeinsame Charakteristiken sind zwar erkennbar (hoher

an Gehalt Glu, Asp), eine weitergehende Ähnlichkeit läßt sich daraus noch nicht ableiten. Auch der direkte Vergleich der Größen der Proteine ist problematisch, da die gefundenen Translationsprodukte im Organismus posttranslational modifiziert werden konnten und deshalb nicht direkt vergleichbar wären.

In einer Reihe elastischer Gewebe lassen sich zwei Glykoproteine (G1 und G2) nachweisen, deren Längen mit den p550-Produkten gut übereinstimmen. Beide treten annähernd in äquimolaren Mengen auf, was ebenfalls für die p550-Proteine zutrifft. G1 und G2 sind jedoch im Gegensatz zum 160 kd Protein nicht durch bakterielle Kollagenase spaltbar. Dieses Argument allein ist noch kein Grund, eine Ähnlichkeit des p550-Proteins mit G1/G2 auszuschließen, da für die 130 und 140 kd Proteine bisher nicht gezeigt wurde, ob sie ebenfalls mit Kollagenase spaltbar sind. Die Unterschiede in der Aminosäurezusammensetzung geben aus den bereits vorher diskutierten Gründen keinen zusätzlichen Hinweis. Interessanterweise finden sich große Mengen dieses Proteins ebenfalls in den Sehnen, in denen auch die mRNA des p550-Gens in großer Menge vorhanden ist.

Ein anderer 'Kandidat' für das p550-Produkt ist das PG-Lt (25), das in Proteingröße, Gewebeverteilung und Kollagenase-spaltbarkeit vergleichbare Eigenschaften zeigt. Aufgebaut ist dieses Protein aus zwei verschiedenen Untereinheiten, die in äquimolaren Mengen vorliegen. Die Aminosäurezusammensetzung weist jedoch einen sehr hohen Gly-Gehalt auf, der im p550-Protein nicht vorhanden ist. Ob dieser Unterschied alleine durch eine ungleiche Aminosäureverteilung in einzelnen Proteindomänen erklärt werden kann, ist fraglich.

Auch in verschiedenen Geweben des Rindes lassen sich Proteine mit vergleichbaren Eigenschaften nachweisen. Ob eines der diskutierten Proteine identisch mit dem p550-Protein ist, muß derzeit noch offen bleiben. Die Wahrscheinlichkeit ist jedoch recht groß, daß das p550-Genprodukt dieser Gruppe von Glykoproteinen zuzurechnen ist, die teilweise in erheblichen Mengen in der Matrix vorliegen

und eine ähnliche Gewebeverteilung zeigen. Erst die Verwendung spezifischer Antikörper wird eine endgültige Zuordnung des p550-Proteins zu bestimmten Matrixkomponenten ermöglichen.

2. Isolation des p550-Gens

Ungleiche Verteilung von Partialfragmenten in den Genbanken

Eine genauere Beschreibung des Expressionsmusters des p550-Gens ist erst sinnvoll im Zusammenhang mit der Struktur des Gens. Aus diesem Grund wurden Genbanken konstruiert und das p550-Gen daraus isoliert. Die Genbanken wurden durch die Klonierung von 15-20 kb langen Eco RI Partialfragmenten in einem λ -Vektor (λ Ch 4A) erzeugt (s.C.2.). Diese Art der Konstruktion fuhrte offensichtlich zu einer ungleichmäßigen Verteilung von Partialfragmenten. Von den insgesamt 9 isolierten genomischen Klonen, die Teile des p550-Gens enthielten (s. Abb. 24), umfassen 8 Klone den gleichen 5'-terminalen Bereich. Nur ein einziger Klon (λ St 550-3) enthält zusätzliche Teile der 3'-terminalen Region. Eine solche Überrepräsentation kann zwei Ursachen haben. Erstens ist denkbar, daß nicht alle Eco RI Stellen mit der gleichen Wahrscheinlichkeit gespalten wurden; nach der Spaltung liegen damit hauptsächlich Fragmente vor, die bevorzugt spaltbare Eco RI Stellen an den Enden aufweisen. Wahrscheinlicher ist jedoch eine zweite Erklärung: die Anordnung der Eco RI Stellen im Bereich des p550-Gens macht die Klonierung einzelner Fragmente unmöglich. Da die Länge der klonierten Fragmente etwa 15-20 kb betrug, müssen alle Partialfragmente, die für eine repräsentative Genbank nötig sind, in diesem Längenbereich zu finden sein. In 'Walking up'-Experimenten gelang es nie, überlappende Klone zu isolieren, die zusätzliche Sequenzen enthielten (s.C.4.3.). Falls

beispielsweise am 5'-Ende des p550 Gens ein Eco RI- Fragment von mehr als 14,1 kb Länge vorliegt, würde ein Partialfragment, das daneben das 8,1 kb Eco RI Fragment enthielte, die maximal klonierbare Insertlänge von 22,2 kb (93) überschreiten. Ein solches Partialfragment wäre damit niemals in der verwendeten Genbank vertreten. Zur Isolierung dieser Region müssen deshalb Fragmente kloniert werden, die mit anderen Restriktionsenzymen erzeugt wurden.

Struktur des p550-Gens

Trotz dieser Limitation in der Vollständigkeit der Genbank war es möglich, etwa 28 kb des p550-Gens zu isolieren. Aus der Analyse der genomischen Klone und dem Vergleich mit dem kodierenden Bereich im Genom konnten wesentliche Eigenschaften des p550-Gens abgeleitet werden :

- (1) Es besteht kein Unterschied in der Genstruktur verschiedener Gewebe.
- (2) Das Gen liegt wahrscheinlich nur einmal im Genom vor.
- (3) Das Gen ist mindestens 25 kb lang und enthält eine große Anzahl von Exons.

Zwei Genbanken waren konstruiert worden, ausgehend von DNA aus Sterna und dedifferenzierten Chondrocyten. Einer der Gründe für dieses Vorgehen war die theoretische Möglichkeit von Unterschieden in der Struktur des Gens in unterschiedlichen Geweben. Durch die Isolierung des Gens aus der DNA eines exprimierenden Gewebes (Sternum) konnte man deshalb sicher sein, das Gen in einer authentischen und funktionellen Form zu isolieren. Auszuschließen waren solche Differenzen durch zwei Befunde: erstens wiesen verschiedene Gewebe keine erkennbaren Unterschiede in ihrem Restriktionsmuster innerhalb des kodierten p550-Genbereichs auf (s.C.3.6.) und zweitens war auch der isolierte genomische Klon (λ De 550-1)

aus der Genbank der dedifferenzierten Chondrocyten identisch mit den Klonen, die aus der 'Sterna'-Genbank isoliert worden waren (s.C.4.4.). Größere Rearrangements, Deletionen, Insertionen etc. des p550-Gens in den untersuchten Geweben konnten damit nicht nachgewiesen werden.

Auf die Anzahl der vorhandenen p550-Gene im Genom kann aus den Ergebnissen der Hybridisierung der klonierten Region mit genomischer DNA geschlossen werden (s.C.3.6.). Tritt ein Gen nur einmal in einem haploiden Chromosomensatz auf, findet sich zwangsläufig ein eindeutiges Restriktionsmuster. Liegen jedoch mehrere Gene vor, können sie sich in ihrem Restriktionsmuster unterscheiden, zumindest in nichtkodierenden Randbereichen. Da im Falle des p550-Gens ein eindeutiges Hybridisierungsmuster auftritt, sollte dieses Gen im Genom nur in einer Kopie vorliegen. Diese Interpretation ist zwar naheliegend, aber durch dieses Ergebnis allein noch nicht bewiesen. Ein gleiches Resultat wäre zu erwarten, wenn vollständig identische Kopien vorhanden wären oder die klonierte Region keine divergierenden Randfragmente enthielte.

Aus der Eindeutigkeit des Restriktionsmusters läßt sich noch ein weiterer Schluß ziehen: innerhalb eines Bereichs von 28 kb gibt es keine mehrmals im Genom auftretenden Sequenzen. Im Genom höherer Eukaryonten finden sich häufig repetitive Elemente, die bis zu 80% der gesamten DNA repräsentieren können (88). Beispielsweise enthalten 95% aller 15-20 kb langen Fragmente humaner DNA repetitive Sequenzen, die der 'Alu I'-Familie zugerechnet werden (124). Das Genom des Huhns enthält demgegenüber nur einen sehr geringen Anteil (12%) repetitiver Sequenzen (125). Anscheinend ist keine dieser Sequenzen aber im Bereich des p550-Gens vorhanden.

3. Exonstruktur des p550 Gens

Durch Restriktionskartierung, Southernblotting und Sequenzierung ließ sich die Sequenz des cDNA Klons p550 im Gen wiederfinden und ist dort auf 5 Exons verteilt (s.C.5.2.). Mit Ausnahme des Exons 4 (219 bp) weisen die Längen aller anderen Exons keine großen Unterschiede auf (105-132 bp). Die Möglichkeit, daß das Exon 4 ebenfalls in Form zweier ähnlich langer, benachbarter Exons vorliegt, kann derzeit nicht ausgeschlossen werden, da dieser Genombereich nicht sequenziert wurde. Eine andere Eigenschaft weisen alle Exons gemeinsam auf: die Übergänge zu angrenzenden Introns liegen immer zwischen den Codons, nie innerhalb eines Triplets, wie es bei den Genen der meisten Proteine der Fall ist (126).

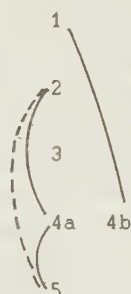
Diese beiden Eigenschaften : ähnliche Längen der Exons und Lage der Exon-Intron Übergänge zwischen den Codons, könnten auf eine mögliche Verwandtschaft der Exons untereinander hindeuten. Denkbar ist eine Entstehung aus gemeinsamen Grundelementen, was sich bei einem Vergleich der Proteinsequenzen der einzelnen Exons an Sequenzhomologien erkennen lassen sollte. Aus diesem Grund wurden die Proteinsequenzen der einzelnen Exons untereinander verglichen. Das verwendete Computerprogramm wertete dabei nicht die Übereinstimmung einzelner Aminosäuren, sondern die Ähnlichkeit auf der Ebene der Codons (72). Die in Tabelle 20 angegebenen SD-Einheiten geben die Abweichungen der gefundenen Ähnlichkeiten zweier Exons an gegenüber einem Vergleich der statistisch gemischten Exonsequenzen. Je größer die Abweichung (SD-Einheit), desto wahrscheinlicher ist die Verwandtschaft zweier Sequenzen.

Eine signifikante Ähnlichkeit zeigte sich nur beim Vergleich der Exons 1 mit 4 (SD=3.5) und 4 mit 5 (SD=4.0). Geringere Werte fanden sich beim Vergleich von 2 mit 4

(SD=2.3) und 2 mit 5 (SD=1.7). Interessanterweise erstreckt sich der homologe Bereich von Exon 1 im Exon 4 nur auf die zweite Hälfte (4b) des Exons, während die andere Hälfte (4a) einen homologen Bereich zu den Exons 2 und 5 aufweist. Damit lassen sich 'Verwandtschaftsbeziehungen' der Exons untereinander angeben. Die gefundene Zweiteilung des Exons 4 gibt vielleicht einen Hinweis darauf, daß dieses Exon tatsächlich in zwei Teilen vorliegt oder zumindest aus zwei unterschiedlich strukturierten Elementen aufgebaut ist.

Tabelle 20 : Vergleich der Exons im Bereich der p550-Sequenz (SD-Einheiten)

| Exon | 1 | 2 | 3 | 4 | 5 |
|------|---|-----|------|------------|------------|
| 1 | | 0.4 | -0.6 | <u>3.5</u> | -0.8 |
| 2 | | | -0.2 | 2.3 | 1.7 |
| 3 | | | | -1.3 | -1.1 |
| 4 | | | | | <u>4.0</u> |



Der Rückschluß auf zwei Ausgangsexons A und B, von denen sich die 5 Exons ableiten lassen (A zu 1, 4b; B zu 2, 3, 4a, 5), ist zwar denkbar, aber sehr hypothetisch. Eine genauere Aussage ist vielleicht möglich, wenn eine größere Anzahl von Exons verglichen werden kann. Eine solche Entstehung wäre deshalb interessant, da das p550-Protein keine auffällige Repetition der Proteinsequenz aufweist. Vergleicht man die DNA- und Proteinsequenzen einzelner Exons untereinander, findet man dennoch auffällige Sequenzhomologien. Im Fall der Kollagengene, bei denen die Amplifikation eines kurzen (54 bp langen) Elements angenommen wird (53), sind solche Homologien deutlich erkennbar. Beim Fibronektin-Gen scheint ein vergleichbarer Fall (etwa 150 bp lange Grundelemente) vorzuliegen (128).

Größe des p550-Gens

Anfang und Ende des Gens konnten bisher noch nicht eindeutig festgelegt werden (s.C.4.2.) und liegen möglicherweise außerhalb des bisher erfaßten Genbereichs. Die Mindestlänge des p550-Gens muß mit etwa 25 kb angenommen werden, wenn man den theoretisch denkbaren Fall außer Acht läßt, daß zwei eng benachbarte Gene ein ähnliches Muster der Expression (mRNA-Längen, Gewebespezifität) zeigten. Vergleicht man die Länge des Gens (>25 kb) mit den Längen der kodierten mRNAs (4.3 kb, 6.5 kb), ergibt sich, daß weniger als 20% der genomischen Sequenzen sich in der mRNA wiederfinden. Diese Relation entspricht etwa dem Verhältnis, das auch im Bereich der p550 cDNA-Sequenz vorliegt (0.55 kb : 2.6 kb). Das Gen muß damit sehr viele und/oder sehr lange Introns enthalten. Nimmt man die Exonlängen des cDNA Bereichs als repräsentativ an (etwa 100-200bp), sollte das gesamte Gen etwa 40-50 Exons aufweisen. Ob tatsächlich eine so starke Fragmentierung vorliegt, konnte noch nicht gezeigt werden, da es bisher nur in wenigen Fällen gelang, einen Teil der Exons in R-Loop Experimenten sichtbar zu machen. Eine ähnliche Anzahl von Exons fand sich nur bei wenigen Genen: den verschiedenen Kollagen-Genen (53), beim Vitellogenin-Gen in *X. laevis* (127) und dem Fibronektin-Gen (128).

4. Hypothetische Promoterregion des p550-Gens

Auf der Suche nach dem 5'-Ende des Gens und damit dem Promoter des p550-Gens wurde der am weitesten 5'-terminal gelegene klonierte Bereich genauer untersucht (s.C.6.4.). Ob in dieser Region bereits das 5'-Ende des Gens enthalten ist, läßt sich derzeit nicht entscheiden, da noch nicht sicher ist, ob das gesamte Gen in den isolierten genomischen Klonen vollständig repräsentiert ist. Auffällig ist dennoch die Häufung ungewöhnlicher Sequenzen bzw. Strukturen am 5'-Ende

des charakterisierten Bereichs :

- (1) Fehlende Methylierung von mindestens 3 CCGG-Sequenzen
- (2) Mehrere mögliche Promotersequenzen
- (3) Region mit einem GC-Gehalt von über 70%
- (4) Mögliche repetitive Sequenz

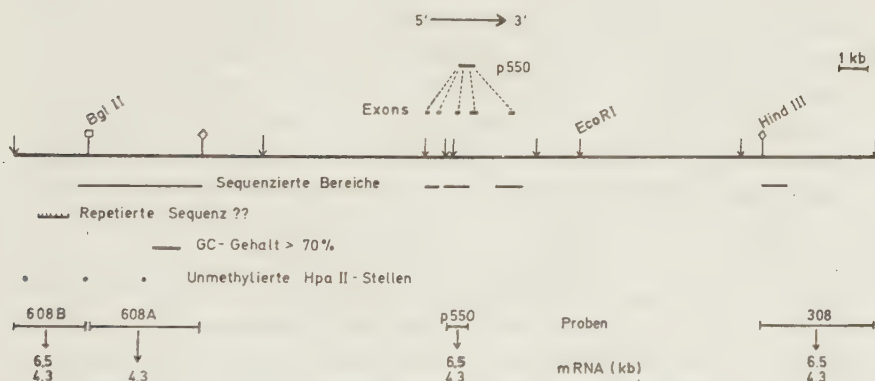


Abb. 49 : Struktur des p550-Gens

In vielen Fällen konnte ein Zusammenhang zwischen dem Grad der Methylierung von CG-Sequenzen und der Aktivität von Genen nachgewiesen werden (98). Allgemein scheint das Ausmaß der Methylierung in einem umgekehrten Verhältnis zur Aktivität eines Gens zu stehen. Detailliertere Untersuchungen zeigten, daß entweder die Methylierung definierter Positionen entscheidend ist (129) oder die Methylierung einen notwendigen, aber nicht hinreichenden Faktor für die Kontrolle der Genaktivität darstellt (99). In anderen Fällen konnte aber ein solcher Zusammenhang von Methylierung und Genaktivität nicht nachgewiesen werden (130). Dennoch

scheint ganz allgemein die Demethylierung bestimmter Abschnitte oder Positionen eine Funktion bei der Markierung transkriptionell aktiver Bereiche zu haben (131). Das Auftreten unmethylierter Sequenzen deutet damit auf die mögliche Nähe einer Promoterregion hin.

Innerhalb des p550-Gens gibt es keine Unterschiede im Muster der Methylierung der DNA aus Sterna, dedifferenzierten Chondrocyten und Kalvarien (s.C.3.7.). Nichtmethylierte Sequenzen wurden ausschließlich im 5'-terminalen Teil des Gens gefunden (608A, 608B).

Der Nachweis der Methylierung war nur auf einen Teil der methylierbaren CG-Sequenzen beschränkt, da bei der Spaltung mit Hpa II/Msp I nur die Sequenzfolge CCGG erfaßt und außerdem entstehende, kleine Fragmente (<400 bp) nicht mehr nachgewiesen werden konnten. Dennoch ist die fehlende Methylierung typisch für diesen Teil des Gens, da alle anderen erfaßten CCGG-Sequenzen in methylierter Form vorlagen. Gerade in diesem Abschnitt, in der Umgebung der Bgl II Stelle (s. Abb. 50), fanden sich jedoch mehrere potentielle Promotersequenzen(s.C.6.4.). bestehend aus 'TATA-Box' und 'CAT-Box' (104). Inwieweit diesen Sequenzen eine funktionelle Bedeutung zukommt, bleibt noch zu klären. Ihre Lage innerhalb des unmethylierten Abschnitts könnte jedoch auf eine aktive Rolle bei der Transkription hindeuten.

In der gleichen Region des Gens fanden sich noch zusätzliche ungewöhnliche Sequenzelemente. Die auffälligste Eigenschaft der Sequenz im Bereich des Fragments 608A ist ein extrem GC-reicher Abschnitt, dessen GC-Gehalt teilweise 90% erreicht (s.C.6.4.). Der durchschnittliche GC-Gehalt liegt immer über 70% und macht es deshalb wenig wahrscheinlich, daß sich hier lange protein-kodierende Sequenzen finden. Diese Sequenzfolge enthält, wie aus Computeranalysen abgeleitet werden konnte, eine große Anzahl von 'Inverted Repeats', die potentiell zur Ausbildung von Sekundärstrukturen führen könnten. Die theoretisch denkbaren 'Stemloops' sollten sehr stabil sein, da sie größtenteils durch GC-Paare gebildet werden. Falls solche Strukturen auch im Genom auf-

treten, könnte ihnen eine Signalfunktion bei der Regulation der Genexpression zukommen.

Ein zweites auffallendes Element wurde nur aus den Daten der Restriktionskartierung abgeleitet, bisher aber noch nicht genauer untersucht (s. C.6.1.) : bei der Spaltung mit Hae III fand sich eine regelmäßige Anordnung von Spaltstellen, deren Abstand einem Vielfachen eines 100 bp Fragments entsprach (100 - 200 - 400 bp). Obwohl die Sequenzierung dieser Region noch aussteht, weist diese Beobachtung auf das Vorliegen eines repetitiven Elements hin, das durch die Amplifikation eines 100 bp Elements entstanden sein könnte. Falls diese Annahme richtig sein sollte, kann diese Sequenzfolge jedoch ausschließlich im p550-Gen vorhanden sein, da nichts auf eine häufig im Genom wiederkehrende Sequenz innerhalb des p550-Gens hindeutete. Inwieweit einem solchen Element eine Funktion bei der Regulation der Genaktivität (132) zukommt, muß beim gegenwärtigen Stand der Untersuchungen offen bleiben.

Die Länge der Repetition, ein Vielfaches eines 100 bp Elements, ist vergleichbar mit der Länge einer Windung der DNA um ein Nukleosom im Chromatin (133). Etwa 195 bp umfaßt der gesamte Bereich eines Nukleosoms auf der DNA. Diese Ähnlichkeit verleitet zu der Spekulation, daß die 100 bp Repetition etwas mit der Positionierung der Nukleosomen zu tun haben und dadurch ebenfalls einen Einfluß auf die transkriptionelle Aktivität des Gens ausüben könnte. Möglichen Veränderungen in der Chromatinstruktur und der Lage einzelner Nukleosomen wurde bereits eine wichtige Rolle bei der Steuerung der Genexpression zugesprochen (96).

Keiner der gefundenen Sequenzen (hypothetische Promotersequenzen, nichtmethylierte CCGG-Sequenzen) oder Strukturen (GC-Bereich, Repetition) kann derzeit eine aktive Funktion bei der Regulation und Initiation der Transkription zugeordnet werden. Dennoch bleibt auffällig, daß in einem kleinen Abschnitt des Gens (608A und 608B) alle diese Elemente gemeinsam auftreten. Dies läßt die Vermutung zu, daß ihnen eine Rolle bei der Regulation und Initiation der Transkrip-

tion zukommen könnte.

5. Transkription des p550-Gens

Die Transkription des p550-Gens führt zur Synthese von zwei mRNA-Species mit unterschiedlichen Längen, 4.3 und 6.5 kb (s.C.1.4.). Diese beiden mRNAs fanden sich bei der Hybridisierung mit einer Reihe von Fragmenten des klonierten Genbereichs des p550-Gens, mit Ausnahme des Fragments 608A (s.C.4.2.). Sequenzen dieses Fragments treten nur in der 4.3 kb mRNA auf (s. Abb. 49). Aufgrund dieses Verteilungsmusters kann der einfachste Fall bereits ausgeschlossen werden, daß die 6.5 kb mRNA eine Vorstufe der 4.3 kb mRNA darstellt, da es nicht vorstellbar ist, daß in einer Vorläuferform Teile der finalen mRNA (608A Sequenzen in der 4.3 kb mRNA) nicht enthalten sein sollen. Aufgrund dieser Tatsache lassen sich die Ergebnisse nur auf zwei Weisen erklären :

- (1) Differentielles Splicing eines primären Transkripts
- (2) Spezifisches Splicing mehrerer primärer Transkripte

Entsteht bei der Transkription nur ein primäres Transkript, muß angenommen werden, daß die Exons im Bereich des 608A Fragments im Falle der 6.5 kb mRNA durch Splicing entfernt werden, nicht jedoch bei der 4.3 kb mRNA. Bei einem solchen 'differentiellen Splicing' konnten mRNAs entstehen, die für identische (136,137) oder auch unterschiedliche Proteine (138,139,140) kodieren. Im Falle des p550-Gens kann man annehmen, daß im größten Teil des Gens das Splicing nach dem gleichen Muster erfolgt, da keine Unterschiede zwischen verschiedenen Hybridisierungsproben (608B, p550, pSt308: s.Abb. 49) im Muster der hybridisierenden mRNAs zu finden sind. Dies scheint sich auch für alle anderen Bereiche des p550 Gens zeigen zu lassen (75).

Falls das differenzielle Splicing auf das 5'-Ende des primären Transkripts beschränkt ist, kann die Längendifferenz der beiden mRNA Species nur durch die Annahme zusätzlicher Unterschiede am 3'-Ende erklärt werden. Differenzielle Termination der Transkription oder die Verwendung verschiedener Polyadenylierungsstellen konnten erklären, warum eine mRNA (6.5 kb) länger ist, obwohl sie einzelne Exons nicht mehr enthält. Solche Unterschiede am 3'-Ende finden sich in vielen Genen (134,135), reichen alleine aber nicht aus zur Erklärung der Längendifferenz von 2.2 kb.

In einigen Fällen konnte gezeigt werden, daß die Transkription an zwei (141,142) oder mehr (143,144) alternativen Stellen beginnt und zur Bildung unterschiedlicher mRNAs führt, die jedoch in den meisten Fällen nicht in verschiedene Proteine übersetzt werden. Erst in einem Fall, einem Myosin-Gen des Huhns (145), konnte gezeigt werden, daß die zwei entstehenden primären Transkripte in unterschiedlicher Weise prozessiert werden und in zwei Proteine übersetzt werden. Die Proteine unterscheiden sich in diesem Fall nur am N-terminalen Ende, da das Splicing in großen Teilen der beiden mRNAs in identischer Weise abläuft.

Besonders der Fall einer differentiellen Initiation, kombiniert mit dem definierten Splicing der unterschiedlichen primären Transkripte bietet eine überzeugende Möglichkeit, sowohl das mRNA Muster und die gefundene Verteilung von Exons (im Bereich 608A und 608B), als auch die Entstehung von drei verschiedenen Proteinen zu erklären. Die einfachste Annahme ist dabei, daß 3 Promotoren existieren, 3 primäre Transkripte gebildet werden, diese in definierter Weise prozessiert werden und zur Synthese von 3 Proteinen führen. Diese Proteine sollten sich dann nur im N-terminalen Bereich unterscheiden, falls Splicing und Leseraster im größten Teil des Gens identisch sind.

Für die Verwendung einer der hypothetischen Promotersequenzen in der Umgebung der Bgl II Stelle (s.C.6.4.) gibt es derzeit zwar keinen Beweis, für den vorgeschlagenen Me-

chanismus würden sie sich jedoch geradezu anbieten (P2 und P3 in Abb. 50 = 'Promoter A, B oder C in Tab. 13/ S.108) Bei der Initiation 'vor' der Bgl II Stelle, damit vor dem 608A Bereich, ist es möglich, daß alle Exons dieses Bereichs durch Splicing entfernt werden. Wird jedoch 'nach' der Bgl II Stelle, innerhalb des 608A-Bereichs initiiert, können diese Exons nicht mehr entfernt werden, da die nötigen Übergangssequenzen nicht mehr vorhanden sind. Diese mRNA sollte also alle Exons des 608A Fragments enthalten. In Abb.50 ist dies schematisch dargestellt. Solange aber nicht bestimmt worden ist, ob eine dieser Initiationsstellen tatsächlich benutzt wird, bleibt diese Beschreibung der Transkription des p550-Gens hypothetisch.

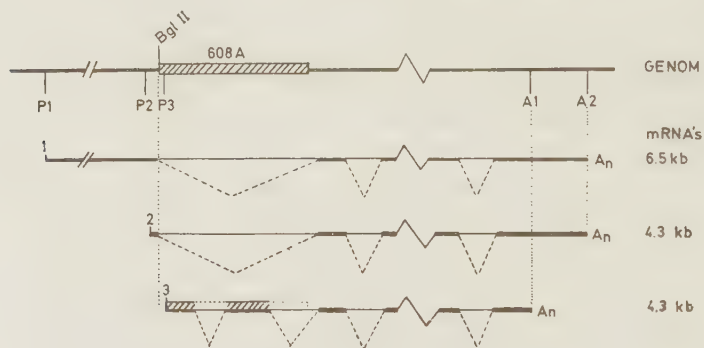


Abb.50 : Hypothetisches Schema von Transkription und Splicing des p550-Gens

P1,P2,P3 : Hypothetische Promotoren

: Intron-/ Exonbereiche auf der mRNA

: Identisches Splicing

Das p550-Gen scheint damit neben dem komplexen Aufbau (große Anzahl von Exons) auch noch in einer komplizierten Weise exprimiert zu werden. Zudem ist das p550-Gen eines der ersten Gene, das eine Komponente der extrazellulären Matrix kodiert, die nicht einem der Kollagene (52) oder dem Fibro-

nektin (128) entspricht. Aus diesen beiden Gründen kann die Untersuchung dieses Gens in nächster Zukunft weitreichende Aufschlüsse über Struktur und Expression dieser Gruppe von Komponenten geben, obwohl die Identität dieses Gens mit einer bestimmten Komponente der Matrix noch nicht gezeigt werden konnte, und selbst die Zuordnung zu dieser Gruppe von Proteinen durch indirekte Hinweise erfolgte (Verteilungsmuster der Expression, Größe und Art der Proteine).

6. Untersuchung weiterer cDNA Klone

Neben dem Klon p550 wurden noch zwei weitere cDNA Klone untersucht (p1300 und p2200), die eine ähnliche gewebespezifische Verteilung der Expression aufweisen (s.C.7.). Alle drei cDNA Klone repräsentieren dabei unabhängige Komponenten 'elastischer' Bindegewebe, deren Expression offensichtlich in vergleichbarer Weise reguliert wird. Dabei ähnelt der Klon p1300 sehr stark dem Klon p550, was die Länge der mRNA (4.3 kb) und deren Verteilung in verschiedenen Geweben betrifft. Dennoch entspricht dieser Klon einem unabhängigen Gen, da nichts darauf hindeutet, daß Sequenzen von p1300 im Bereich des p550 Gens zu finden sind. Der durchgehende Leseraster des Klons, der durch die Bestimmung der Orientierung der Sequenz auf der mRNA sehr wahrscheinlich gemacht wurde, gibt auch in diesem Fall keinen direkten Hinweis auf die Identität des Klons. Keine der bisher bekannten Sequenzen zeigt eine Ähnlichkeit mit der cDNA Sequenz. Aus der Aminosäurezusammensetzung, die etwa 30% der Gesamtsequenz entspricht (1.3 kb cDNA : 4.3 kb mRNA), läßt sich zumindest auf die Art des Proteins schließen, da besonders der hohe Gehalt an Thr, Ser und Cys auffällt. Diese Zusammensetzung deutet möglicherweise auf ein sehr stark quervernetztes (Cys) Proteoglykan hin, das eine große Anzahl von Glykosaminoglykanseitenketten gebunden (an Ser und Thr) enthält (18). Bisher gelang es noch nicht, durch Hybridselektion die Größe

des kodierten Proteins zu bestimmen, sodaß eine genauere Zuordnung zu einer bestimmten Komponente derzeit nicht möglich ist.

Der dritte untersuchte cDNA-Klon (p2200) weist zwar ein sehr ähnliches gewebespezifisches Expressionsmuster auf, wie es bereits bei p550 und p1300 gefunden wurde, die kodierte mRNA unterscheidet sich jedoch sehr deutlich in ihrer Länge (7.8 und 11 kb). Die Kodierungskapazität einer solchen mRNA reicht aus für ein Protein von mehr als 200 kd. Dies war auch der Grund, diese cDNA mit dem Proteinanteil eines Proteoglykans in Verbindung zu bringen (61). Die Sequenz dieses cDNA Klons wies jedoch keinen durchgehenden Leseraster auf. Zudem liegt die Sequenz unverändert im Genom vor, enthält also keine trennenden Intronsequenzen (s.C.7.4.). Dies läßt sich auf zwei Weisen erklären :

- (1) Der cDNA Klon p2200 entspricht einer, nicht vollständig processierten Zwischenstufe des Splicings.
- (2) Die cDNA enthält terminale Sequenzen der mRNA, die nicht oder nur teilweise für Protein kodieren.*

Für den ersten Fall spricht die Tatsache, daß sich in der Sequenz von p2200 potentielle Exon-Intron Übergänge nachweisen lassen. Der zweite Fall setzt dagegen voraus, daß die gesamte cDNA Sequenz einem einzelnen Exon zugerechnet werden muß. Welche der beiden Möglichkeiten richtig ist, läßt sich derzeit noch nicht sagen. Damit läßt sich natürlich auch nicht entscheiden, ob die 7.8 kb mRNA die finale translatierbare Form darstellt, da nicht auszuschließen ist, daß die cDNA Sequenz als Ganzes der Sequenz eines Introns entspricht. In diesem wenig wahrscheinlichen Fall wären nur einzelne Zwischenstufen des Splicings mit p2200 als Hybridisierungsprobe nachweisbar.

Die Zuordnung zu einer bestimmten Matrixkomponente steht auch im Falle dieses cDNA Klons noch aus.

7. Verwendung spezifischer cDNA Klone zur Untersuchung der Expression von Matrixkomponenten

Die drei untersuchten cDNA Klone entsprechen unterschiedlichen, in ähnlicher Weise regulierten Matrixkomponenten, deren Expression nur in bestimmten Bindegeweben erfolgt (Sternum, Kalvarien, Sehne). Trotz der noch fehlenden Identifizierung der einzelnen kodierten Proteine bieten diese Klone die Möglichkeit, molekulare Regulationsvorgänge bei der koordinierten Expression unabhängiger Gene zu untersuchen. Damit kann ein Einblick in grundlegende Mechanismen bei der Differenzierung von Geweben gewonnen werden.

E. ZUSAMMENFASSUNG

Aufgabe dieser Dissertation war es, durch die Charakterisierung spezifischer cDNA-Klonen und deren homologen Genen ein genaueres Bild von Expression und Struktur einzelner Matrixkomponenten des Knorpels zu erhalten. Drei unterschiedliche cDNA Klone (p550, p2200 und p1300) wurden in dieser Arbeit untersucht. Aus den Ergebnissen lassen sich folgende Rückschlüsse auf die Identität der Genprodukte und deren gewebespezifischer Expression ableiten :

- (1) Die drei cDNA-Klone entsprechen Sequenzen dreier unterschiedlicher mRNAs und damit dreier unabhängiger Gene. Die Längen der homologen mRNAs betragen :
p550 = 4.3 + 6.5 kb
p2200 = 7.8 + 11 kb
p1300 = 4.3 kb.
- (2) Die Expression der spezifischen mRNAs erfolgt in koordinierter Weise nicht nur im Knorpel des Sternums, sondern auch in den Sehnen des Hühnerembryos. In den Kalvarien läßt sich die Expression nur in frühen Stadien der Embryonalentwicklung nachweisen.
- (3) Im Falle des cDNA-Klones p550 konnte die Größe der kodierten Proteine durch 'Hybrid-Selektion' und In Vitro Translation der spezifisch angereicherten mRNA bestimmt werden. In drei unterschiedliche Proteine läßt sich die p550 entsprechende mRNA übersetzen ; 160, 140 und 130 kd.
- (4) Die abgeleiteten Proteinsequenzen der cDNA-Klone und das Verteilungsmuster der Expression der entsprechenden Gene entsprechen keinen bisher beschriebenen Komponenten der Knorpelmatrix.

Die Untersuchung dieser cDNA-Klone bietet damit eine Möglichkeit, Komponenten der Knorpelmatrix zu erfassen, die mit proteinchemischen Methoden kaum analysierbar sind. Um auch einen Zugang zu den Genen dieser Komponenten zu erhalten, wurden repräsentative Genbanken konstruiert, ausgehend von der DNA exprimierender Gewebe (Sternum) und nicht exprimierender Zellen (Dedifferenzierte Chondrocyten). In dieser Arbeit wurde die Isolierung und detaillierte Untersuchung eines dieser Gene ('p550-Gen') beschrieben. Dabei ergaben sich für dieses Gen folgende wesentlichen Merkmale :

- (1) Das p550-Gen umfaßt im Genom einen Bereich von mindestens 25 kb und ist in eine größere Anzahl von Exons unterteilt (etwa 20% des gesamten Gens). Der Bereich der cDNA kodierenden Exons wurde im Detail beschrieben.
- (2) Am 5'-Ende des Gens wurde ein Bereich genauer untersucht, der wahrscheinlich die Promoterregion des Gens enthält und verschiedene Besonderheiten aufweist : nichtmethylierte Hpa II-Stellen, Repetitive Sequenzen und eine extrem GC-reiche Region.
- (3) Die Expression dieses Gens : 1 Gen - 2 mRNAs - 3 Proteine erfolgt wahrscheinlich durch 'differentielle Initiation' an verschiedenen Promotoren und ein definiertes Spleißen der entstehenden primären Transkripte. Drei unterschiedliche mRNAs führen zur Synthese der drei unterschiedlichen Proteine, die sich wahrscheinlich nur am N-Terminus unterscheiden.

Die detailliertere Untersuchung der Regulationsvorgänge dieser Gene ermöglicht damit einen Einblick in die molekularen Prozesse bei der koordinierten Expression unterschiedlicher Gene im Knorpel.

F. ABKÜRZUNGEN

| | |
|-------|--|
| ABM- | Aminobenzoxymethyl- |
| bp | Basenpaare |
| BrdU | 5-Bromdesoxyuridin |
| Ci | Curie |
| dNTP | Desoxynukleotidtriphosphat |
| ddNTP | 2'3'-Didesoxynukleotidtriphosphat |
| DAP | Diaminopimelinsäure |
| DMF | Dimethylformamid |
| DMSO | Dimethylsulfoxid |
| DT | +Diaminopimelinsäure/ Thymin |
| DTT | Dithiothreitol |
| DBM- | Diazobenzoxymethyl- |
| EDTA | Ethylendiaminetetraessigsäure |
| FKS | Fötales Kalberserum |
| kb | Kilobasen = 1000 b bzw. bp |
| kd | Kilodalton |
| Min. | Minuten |
| oD | optische Dichte |
| PEG | Polyethylenglykol |
| pfu | 'Plaque forming units' (Phagentiter) |
| RT | Raumtemperatur |
| SDS | Natriumdodecylsulfat |
| Std. | Stunden |
| U | Enzymeinheiten |
| n.b. | nicht bestimmt |
| rpm | Umdrehungen pro Minute |
| XGal | 5-Brom-4-Chlor-3-Indolyl- β D-Galaktosid |

G. Materialien und Methoden

1. Materialien

Chemikalien

Die meisten Laborchemikalien, Harnstoff und Glycin wurden von Merck und Riedel de Haen bezogen; Lysozym, Thymin, Ampicillin, Diaminopimelinsäure, Spermidin, Putrescin und Ethidiumbromid von Sigma; Guanidiniumhydrochlorid, Phenol, IPTG, CsCl und Saccharose von BRL; Streptomycin, Uridin, TEMED, PEG 6000, Triton X-100, Bisacrylamid, SDS und Amberlit MB-3 von Serva; Ampicillin, Tetracyclin und X-Gal von Boehringer; Mercaptoethanol, DMSO und Acrylamid von Roth; Bind- und Repelsilan von LKB; Dimethylformamid von Baker; Sarkosyl NL-30 war ein Geschenk von Ciba-Geigy; Glyoxal von EGA-Chemie.

Materialien für die Molekularbiologie

Restriktionsenzyme von BRL, Boehringer und PL; Ligase und Marker-DNAs (λ Hind III, ϕ X Hae III) von Biolabs; Alkalische Phosphatase, DNA Polymerase, Klenow Fragment, S1-Nuclease, DNase I, RNase A, M13 primer, tRNA (Hefe) und Desoxynukleotide von Boehringer; Polynukleotidkinase, Didesoxynukleotide, M13 primer und M13 Klonierungs- und Sequenzierungskit von PL; In vitro Translationskits von NEN und BRL; Low melting point Agarose von BRL; Agarose von Seakem; Human placental RNase inhibitor (RNasin) von Amersham; Proteinase K von Merck; Bakterielle Kollagenase war ein Geschenk von I. Oberbauer; Genescreen Transfermembran von NEN; Nitrocellulosefilter und 3MM Papier von Schleicher & Schüll; Sephadex G-50 und Sepharose S-1000 von Pharmacia; Macromould Sequenzierapparatur und Zubehör von LKB; Röntgenfilme

(XAR-5, Ortho-G) und lichtdichte Kassetten mit 'Lanex Regular' Verstärkerfolie von Kodak.

Radioaktive Substanzen

$\alpha^{32}\text{P}$ Desoxynukleotide (3000 Ci/mmol) von Amersham und NEN: $\alpha^{35}\text{S}$ Desoxyadenosintriphosphat (>600 Ci/mmol) und $\gamma^{32}\text{P}$ Adenosintriphosphat (3000 Ci/mmol) von Amersham: ^{35}S Methionin (>800 Ci/mmol) von NEN.

Bakterienstämme/ Phagen

C600 : F^- , thi, thr, leu8, lacY, tonA, supE (147)

HB101 : F^- , hsdS, recA, ara, proA, lacY, galK, supE (147)

DP50supF : F^- , tonA, dapD, lacY, glnV, $\Delta(\text{gal-uvrB})$, supE, supF, ΔthyA , tyrT, gyrA (84)

BHB2688 : N205 recA [rimm, cits, b2, red, Eam, Sam]/ γ (92)

BHB2690 : N205 recA [rimm, cits, b2, red, Dam, Sam]/ γ (92)

JM 83 : ara, $\Delta(\text{lac.pro})$, strA, thi, $\Phi 80\text{dlacZAM}$ (101)

JM 103 : $\Delta(\text{lac.pro})$, thi, strA, supE, $\text{F}'\text{traD}$, lac i^q, lacZAM (65)

Medien für die Zellkultur

MEM Dulbecco und Ham's F10 Medium wurden von LS Labor Service bezogen: Fetales Kalbsserum und Medienzusätze von Seromed.

Medien für die Bakterien- und Phagenkultur

Trypton, Hefeextrakt und Agar wurden von Difco bezogen und die Medien (LB, LBM, M9 Minimal) und Platten nach (147) hergestellt. Ausgehend von einer 100xDT Lösung (1g Diaminopimelinsäure, 0.4g Thymin pro 100 ml) wurde den Medien 1xDT zugesetzt. Zusätzlich verwendete Medien und Lösungen sind bei den einzelnen Methoden beschrieben.

2. Methoden

2.1 Zellkultur

Chondrocyten wurden aus 200-400 Sterna von 17-19 Tage bebrüteten Hühnerembryonen (Hühnerbrüterei Hölzel, Moosburg) isoliert und in Kultur genommen (148). Nach mehrmaligem Waschen der Sterna in PBS wurde mit 0,1% Collagenase, 0,3% Trypsin in MEM Dulbecco Medium für 20 min bei 37°C das Perichondrium abgebaut und danach wieder gewaschen. Die Chondrocyten wurden durch Behandlung mit frischer Lösung für 2 Std freigesetzt, in Ham's F10 Medium mit 10% FKS/ 50 U/ml Penicillin, 50 µg/ml Streptomycin, 50 µg/ml Ascorbat 3 mal gewaschen und pelletiert. Die Chondrocyten wurden dann in einer Zelldichte von 1×10^7 Zellen pro Platte (ϕ 10 cm) plattiert und bei 37°C und 5% CO₂ kultiviert. Das Medium wurde nach einem Tag und danach alle 4-6 Tage gewechselt. Die Zellen wurden für 2 Tage (=diff. Chondrocyten) oder für 5-6 Wochen (=dediff. Chondrocyten) kultiviert.

2.2 DNA Isolierung

DNA Isolierung aus Chondrocyten und Sterna

40 Platten mit konfluenten Chondrocyten (etwa 1×10^8 Zellen) wurden nach dem Abwaschen des Mediums mit PBS mit je 1-1,5 ml Lysepuffer (1% Sarkosyl NL-30, 100 mM EDTA, 100 µg/ml Proteinase K) für 5 min bei 56 °C inkubiert. Die Zellen können danach abgeschabt und die RNA in den etwa 50 ml Zellysat mit 10 µg/ml RNase A (10 min bei 80°C vorbehandelt) bei 56°C in 2 Std abgebaut werden. Nach dreimaliger Extraktion mit Phenol:Chloroform:Isoamylalkohol (24:0,5:0,5 v/v) wurde die wässrige Lösung mit Äther extrahiert und

dieser danach abgedampft. 1,25 g CsCl wurden pro ml Lösung zugegeben und in einem VTi 50 Rotor bei 4°C und 30 krpm für 24 Std zentrifugiert. Nach der Fraktionierung des Gradienten wurden die DNA enthaltenden Fraktionen gepoolt und ausgiebig gegen TE- Puffer dialysiert.

Ausbeute : etwa 1.5 mg DNA/ 10^8 Zellen).

Zur direkten Isolierung von DNA aus Sterna wurden 400 frisch isolierte Sterna aus 17 Tage bebruteten Hühnerembryonen möglichst sauber präpariert und in PBS gewaschen. Das Perichondrium wurde abgebaut (s.1.2.1) und die Sterna erneut mit PBS gewaschen. Durch Behandlung mit Collagenase/ Trypsin Lösung für jeweils 90 min bei 37°C wurden die Chondrocyten freigesetzt und nach der Sedimentation der Sternaresten abgenommen, zentrifugiert und gewaschen. Die vereinigten Chondrocyten aus sechsmaligem Verdau der Sternafragmente wurden danach durch Zugabe von 50 ml Lysepuffer lysiert. Die DNA konnte danach wie vorher beschrieben isoliert werden.

Ausbeute : etwa 3mg DNA/ 400 Sterna (17 Tage alt)

Isolierung von Plasmid DNA

Minilyse: Zur Isolierung kleiner Mengen von Plasmid DNA (149) wurden die Zellen von 2.2 ml Übernachtskultur 1 min bei 12 000g pelletiert und mit 0.4 ml Lysozymlösung (2mg/ml Lysozym in 25 mM Tris, 20 mM CDTA, 50 mM Glucose, pH 8) 30 min bei 0°C lysiert. Die DNA wurde durch Zugabe von 0.8 ml 0.2 N NaOH/ 1% SDS für 5 min auf Eis denaturiert und danach 60 min bei 0°C mit 0.6 ml 3 M KAc neutralisiert. Die dabei ausfallenden Proteine wurden 15 min bei 12000g abzentrifugiert und der Überstand mit 770 µl Isopropanol bei -20°C gefällt. Die DNA wurde nach dem Auflösen in 150mM NaCl, 50mM Tris, pH8 mit Phenol:Chloroform (1:1) extrahiert und nochmals gefällt, mehrmals mit 70% Alkohol gewaschen und in TE Puffer gelöst.

Großansatz: Zur Isolierung großer Mengen Plasmid DNA (pBR und pUC Derivate) wurden 2-5 l M9 Minimalmedium (8.9g NaH_2PO_4 / 3g KH_2PO_4 / 0.5g NaCl / 1g NH_4Cl / 10 mM MgSO_4 / 0.4% Glucose/ 0.5% Casamino-säuren/ 10 μg Thiamin/ 0.1 mM CaCl_2 / 0.1 mg/ml Ampicillin) mit 1% Übernachtskultur angeimpft (147) und nach Erreichen einer $\text{OD}_{575}=0.1$ mit 2g/l Uridin versetzt. 100 mg/l Chloramphenicol wurde bei Erreichen von $\text{OD}=0.6$ zugegeben und noch 15 Std weiter geschüttelt. Auf die Zugabe von Uridin wurde nur verzichtet, wenn statt Minimalmedium LB-Medium verwendet wurde. Die Zellen wurden abzentrifugiert und lysiert (Angaben pro l) :

- a. 15 ml 50 mM Tris/ 25% Sucrose/ pH8
- b. 3 ml 5 mg/ml Lysozym in 25 mM Tris/ pH8 für 5 min/4°C
- c. 1.5 ml 250 mM EDTA/ pH8 für 15 min/4°C
- d. 15 ml 0.3% Triton X-100/ 200 mM EDTA/ 150 mM Tris/ pH8 für 30 min bei 4°C.

Die lysierten Zellen wurden bei 30 krpm (50.2 Ti/ 4°C/ 60 min) zentrifugiert und der Überstand mit 25 μg /l RNase A (20 min bei 80°C vorbehandelt) für 30 min bei RT inkubiert. Zu der viskosen Lösung wird 1g CsCl pro ml und 0.1 mg/ml Ethidiumbromid zugegeben, filtriert und zentrifugiert (VTi 50/ 20°C/ 45 krpm/ 50-60 Std). Die Plasmidbande wurde entnommen und 8 mal mit dem gleichen Volumen Isopropanol (gesättigt mit CsCl) extrahiert und danach ausgiebig gegen TE Puffer dialysiert. Falls nötig wurde durch eine Alkoholfällung weiter konzentriert.

Ausbeute abhängig von Plasmidgröße : 1-5 mg DNA /l Kultur

Falls noch RNA in Plasmidpräparationen enthalten war, wurde nochmals mit RNase A behandelt und die DNA durch ein Kissen aus 1 M NaCl zentrifugiert (147).

Isolierung von Phagen DNA

λ -Phagen : Eine Übernachtskultur von DP50supF in TBMM Medium (10g Trypton/ 2.5g MgSO_4 / 2g Maltose/ 5g NaCl / 1xDT) wurde in M-Lösung (10 mM MgSO_4 / 1xDT) 10 fach konzentriert,

2 ml davon mit 4×10^7 pfu Phagen 10 min bei 30°C inkubiert und damit 2 l LBM Medium (LB + 0.5g/l MgSO_4 / $1 \times \text{DT}$ / 5mM CaCl_2) beimpft. Nach 14 Std Kultur bei 32°C wurden 2 ml Chloroform zugegeben und für 15 min bei niedriger Geschwindigkeit weitergeschüttelt. Je $1\text{ }\mu\text{g/ml}$ DNase I und RNase A wurden zugegeben und nach 40 min bei 0°C wurde die Lösung auf 1 M NaCl gebracht und 20 min zentrifugiert. Die Phagen im klaren Überstand lassen sich danach mit 10% PEG 6000 in 7 Std bei 4°C fällen. Das abzentrifugierte Phagenpellet wurde in TMG (50 mM Tris / $\text{pH} 8$ / 6g NaCl / 2.5g MgSO_4 / 0.1g Gelatine / $1 \times \text{DT}$) dem 100 mM NaCl zugegeben waren suspendiert und 8-10 mal mit Chloroform extrahiert. Die Reinigung der Phagen erfolgte danach in einem CsCl Stufengradienten. Zur Phagenlösung wurde 0.5 g/ml CsCl zugegeben und auf einen Gradienten mit den Dichtestufen 1.7, 1.5, 1.45 g/ml CsCl in 20 mM Tris / 10 mM MgSO_4 / $\text{pH } 7.5$) aufgeladen und 2 Std zentrifugiert (50.2 Ti / 15°C / 27 krpm). Die Phagenbande an der Phasengrenze 1.5-1.45 wurde isoliert. mit CsCl Lösung der Dichte 1.5 gemischt und für 15 Std zentrifugiert. Die saubere Phagenbande wurde abgenommen und nach Zugabe von 0.5 mg/ml Proteinase K gegen TE dialysiert ($3 \times 30\text{ min}$ / 100 faches Vol.). Nach 3 maliger Phenolextraktion wurde nochmals ausgiebig gegen TE dialysiert.

Ausbeute : $0.2 - 1\text{ mg DNA}$ / 2 l Kultur

M13 Phage : Zur Isolierung der einzelstrangigen Phagen DNA wurde eine Übernachtskultur von JM 103 in YT + 0.12 mg/ml Streptomycin 1:100 in frisches Medium verdünnt und 2 Std bei 37°C inkubiert. 2 ml der exponentiell wachsenden Kultur wurden mit M13 infizierten Zellen (Plaques von Platte oder $10\text{ }\mu\text{l}$ einer Glycerinkultur) angeimpft und 6.5 Std bei 37°C geschüttelt. Die Zellen wurden danach 5 min bei 12000 g abzentrifugiert und die Phagen des Überstandes mit $1/10\text{ Vol}$ 20% PEG 6000 / 3M NaCl für 10 min bei 0°C gefällt. Die Phagen lassen sich danach abzentrifugieren, in TES (20 mM Tris / $\text{pH } 7.5$ / 20 mM NaCl / 1 mM EDTA) suspendieren, mit Phenol und danach mit Ather extrahieren. Die DNA wurde mit Alkohol

gefällt und in 10 mM Tris/pH 7.5 aufgenommen.

Ausbeute : 2-10 μg DNA/ 2 ml Kultur

Die doppelsträngige replikative Form von M13 wurde aus dem Zellpellet der Präparation mit der alkalischen Lysemethode, wie vorher beschrieben, isoliert wobei die Mengen den Ausgangsvolumina angepaßt wurden.

Bestimmung der DNA Konzentration

Die DNA Konzentration wurde durch Messung der Absorption bei 260 nm gemessen : 1 oD/ml = 50 $\mu\text{g/ml}$ DNA. Die Reinheit wurde aus dem Verhältnis der Extinktionen bei 260 und 280 nm abgeschätzt (saubere DNA = 1.8-2.0). Für einzelsträngige DNA wurde 1 oD/ml = 40 $\mu\text{g/ml}$ angenommen. Zur Messung kleiner Mengen DNA wurden 1-2 μl DNA Lösung mit dem gleichen Volumen 1 $\mu\text{g/ml}$ Ethidiumbromid gemischt und die Fluoreszenz im UV Licht mit einer Eichreihe bekannter Konzentrationen verglichen.

2.3 Restriktionsspaltungen, Gele, Southernblotting

Restriktionsspaltungen

Die DNA wurde entweder in dem von den Herstellerfirmen angegebenen Puffer oder analog einem vereinfachten Pufferschema (147) gespalten. Für analytische Spaltungen wurden verwendet :

| | | | |
|-----------------|---------------------|--------|---------|
| Plasmide | 0.5-2 μg | 1-10 U | 1-3 Std |
| Phage λ | 2-3 μg | 1-10 U | 1-3 Std |
| genomische DNA | 5-10 μg | >25 U | 15 Std |

Für präparative Spaltungen wurden die Bedingungen den DNA Mengen angepaßt.

Gelelektrophorese und DNA Isolierung

Agarosegele : DNAfragmente wurden in horizontalen Agarosegelen (13x13/19x0.4 cm) bei 4-8 V/cm in 1xTBE = 90 mM Tris-borat/ 2.5 mM EDTA (147) aufgetrennt und nach Färbung in 0.5 µg/ml Ethidiumbromid auf einer UV-Lampe bei 254 nm sichtbar gemacht und photographiert (Polaroidfilm Typ 667/ASA 3000 und Kodak 22A Wratten Filter). Die Längenbestimmung erfolgte durch den Vergleich mit Fragmenten bekannter Länge (λ /Hind III und ϕ X/Hae III Spaltungen). Folgende Agarosekonzentrationen wurden verwendet : 0.5% (2-50 kb), 0.8% (0.5-20 kb), 1% (0.3-10 kb). Für schnelle analytische Trennungen wurden kleine (6x4 cm) Agarosegele verwendet.

Zur Isolierung von Fragmenten aus Agarosegelen wurden 2 versch. Methoden verwendet. Nach () wurde Low melting point Agarose verwendet und die DNA aus der geschmolzenen isolierten DNA Bande mit Phenol extrahiert (s. 2.5). Daneben wurde die DNA häufig durch Elektrophorese der DNA in einen Trog vor der Bande (151), der mit Dialyseschlauch ausgekleidet war, eluiert, Phenol extrahiert und gefällt. Bei beiden Verfahren wurde TAE Puffer (40 mM Tris-Acetat/ 2mM EDTA) verwendet. Ausbeute : 40 - 80 % der eingesetzten Menge

Acrylamidgele : Zur Auftrennung kleiner DNA Fragmente wurden 7% Acrylamidgele verwendet (20x20x0.3 cm) in 1x TBE Puffer verwendet. Nach der Auftrennung bei 10 V/cm und dem Färben in Ethidiumbromid können die DNA Fragmente auf einer UV-Lampe wieder sichtbar gemacht werden. Zur Isolation wurden die Banden ausgeschnitten, homogenisiert und mit 500mM NH₄acetat/ 10 mM Mgacetat/ 1mM EDTA über Nacht bei 37°C extrahiert (64). werden. Nach der Filtration durch silinierte Glaswolle kann die DNA mit Alkohol gefällt werden.

Southernblotting

Nach der Auftrennung der DNA Fragmente in 0.8-1% Agarosegelen wurde das Gel in 10 mM HCl für 10 min partiell depuriniert, 1 Std in 0.5 M NaOH/ 1.5 M NaCl denaturiert und danach 1 Std in 0.5 M Tris/ pH7/ 3M NaCl neutralisiert. Der Transfer auf Genescreen Transfermembranen erfolgte nach (94) einen kapillaren Strom von 20xSSC (3M NaCl/ 0.3 M Nacitrat) durch Gel und Membran in einen Stapel aus 3MM Papier und Papierhandtücher passender Größe. Die Filter wurden danach in 2xSSC gewaschen und nach dem Trocknen 3 Std bei 90°C gebacken.

Elektrotransfer : Aus Agarosegelen wurde die DNA auch durch Elektroblotting auf Genescreen transferiert. Die Gele wurden wie vorher beschrieben depuriniert und denaturiert, danach aber in 25 mM Phosphatpuffer/ pH6.5 neutralisiert. Der Transfer erfolgte im gleichen Puffer bei 0.25A für 1-2 Std und danach 0.7A für 12 Std. Die Filter wurden danach im gleichen Puffer kurz gewaschen, luftgetrocknet und gebacken.

2.4 RNA Isolierung und Northernblotting

RNA Isolierung

RNA wurde nach der Guanidiniumhydrochlorid Methode aus konfluenten Monolayer Zellkulturen (etwa 6×10^7 Zellen) oder aus den entsprechenden Organen (5 g Naßgewicht) embryonaler Hühner extrahiert (150). Dem Homogenisieren in 8M Guanidiniumhydrochlorid folgte eine dreimalige Fällung mit dem halben Alkoholvolumen bei -20°C für 30 min. Der Rückstand wurde in 20 mM EDTA/ pH 7.4 gelöst und mit Chloroform : Isobutanol (4:1) extrahiert. Die wäßrige Phase wurde mit 3 Volumina 4 M Naacetat gemischt und über Nacht bei -15°C gefällt. Das RNA Pellet wurde zweimal mit Alkohol gefällt.

Northernblotting

10 μg RNA wurden in 20 μl 50% DMSO/ 1M Glyoxal/ 10 mM Phosphat/ pH 6.5 für 1 Std bei 55°C denaturiert (70) und nach Zugabe von 5 μl 50% Glycerin/ 10 mM Phosphat/ 0.25% Bromphenolblau/ pH 6.5 in einem 1% Agarosegel bei 4 V/cm aufgetrennt. Der Laufpuffer (10 mM Phosphatpuffer/ pH6.5) wurde während des Laufes dauernd umgewalzt. Nach dem Lauf (Laufstrecke des Bromphenolblau etwa 12 cm) wurde die RNA ohne weitere Behandlung des Gels mit 20xSSC auf Genescreen transferiert. Nach dem 15-20 stündigen Transfer wurde die Membran nicht gewaschen sondern nach dem Trocknen 3 Std bei 90°C gebacken. Zur Längenbestimmung der RNA wurde endmarkierte $\lambda\text{Hind III}$ Marker DNA auf die gleiche Weise behandelt und transferiert. Nach dem Transfer ließ sich die Vollständigkeit des Transfers an der UV Absorption der RNAs auf dem Filter erkennen.

RNA wurde auch durch Elektroblothing transferiert, wobei der Transfer ebenfalls in 10 mM Phosphat/ pH 6.5 bei 0.7A/ 4°C für 15 Std erfolgte.

2.5 Radioaktive Markierung von DNA/RNA und Hybridisierung

Nicktranslation

DNA Fragmente wurden durch Nick Translation (154) radioaktiv markiert. 0.2-1 μg DNA wurden in 50 mM Tris/ pH7.5/ 10mM MgCl_2 / 1mM DTT/ 30 μM unmarkierten dNTPs und 100 μCi $\alpha^{32}\text{P}$ -dATP mit 100-200 pg DNase I für 15 min behandelt. 5-10 U DNA Polymerase I wurden zugegeben und weitere 2 Std bei 15°C inkubiert. Die Reaktion wurde mit 50 mM EDTA gestöppt und nicht eingebaute Nukleotide durch Chromatographie an Sephadex G-50 (0.5x15 cm Säule) abgetrennt. In einzelnen

Fällen wurden auch mehrere markierte Nukleotide verwendet.

Ausbeute : $0.2-1 \times 10^8$ cpm/ μ g DNA

Strangspezifische Markierung von M13 DNA

0.1-0.5 μ g einzelsträngige M13 DNA wurde in 10 μ l 6mM Tris/ pH 7.5/ 6 mM NaCl/ 6mM $MgCl_2$ / 1mM DTT mit 3 ng 'Hybridisierungsprimer' (5'-GTCATAGCTGTTTCCTG-3') Primer für 10 min bei 55°C inkubiert und danach langsam auf RT abgekühlt. Nach der Zugabe von 40-100 μ Ci radioaktiv markiertem Nukleotid und je 50 μ M der restlichen Nukleotide wurde durch Zugabe von 0.5 U Klenow Fragment (Vol=20 μ l) die Reaktion gestartet. Nach 1 Std bei 15°C wurde mit 50 μ M EDTA gestoppt und der Ansatz 2 mal mit Alkohol gefällt und mehrmals gewaschen.

Ausbeute : $0.1-1 \times 10^8$ cpm/ μ g M13 DNA

Hybridisierungen

Filtergebundene DNA und RNA wurden gleichermaßen mit der Sarkosyl Methode (71) hybridisiert. Die Filter wurden in einem Volumen von 10-20 ml 5xSSC/ 1% Sarkosyl NL-30 mindestens für 1 Std bei 65°C in einer verschlossenen Tüte geschüttelt. Nach der Prähybridisierung wurde ein möglichst kleines Volumen frische Hybridisierungslösung (3-5 ml/100 cm²) zu den Filtern gegeben, denaturierte (5 min/ 95°C), nicktranslatierte DNA oder markierte M13 DNA (ohne Denaturierung!) zugegeben und für 15-50 Std weiter inkubiert. Meist wurden etwa $2-10 \times 10^4$ cpm/cm² verwendet. Nach der Hybridisierung wurden die Filter 2 mal 30 min in 3xSSC/ 0.5% Sarkosyl-NL 30 und danach 2 mal 30 min in 3xSSC jeweils bei 65°C gewaschen. Wenn die Filter nochmals hybridisiert werden sollten wurden sie noch leicht feucht in Folie luftdicht verpackt, in allen anderen Fällen wurden die Filter getrocknet. Die Exposition erfolgte, je nach der Intensität der Signale für verschiedene Zeiten in lichtdichten Kassetten mit oder ohne Verstärkerfolien.

Markierung von RNA und Hybridisierung

Sterna RNA wurde zuerst in Agarose-Harnstoff Gelen aufgetrennt und daraus extrahiert (152). Etwa 50 µg Sterna RNA (19d) wurde in 40% Harnstoff und 0.5xTAE Puffer aufgenommen, 2 min bei 95°C denaturiert und in einem vertikalen 1% igen Low melting point Agarose Gel mit 5 M Harnstoff (0.15x20x20 cm) in 1xTAE Puffer bei 5 V/cm und 4°C in 5 Std aufgetrennt. Nach dem Färben des Gels wurden die Bandenbereiche etwa 4.3-5 kb ausgeschnitten und bei 65°C geschmolzen. Das 3 fache Volumen ANE (10mM Naacetat/ 100mM NaCl/ 1mM EDTA/ pH 6.8) + 0.1% SDS wurde zugegeben und nach 10 min bei 65°C mit ANE gesättigtem Phenol extrahiert. Die wäßrige Phase wurde mit Ethanol gefällt und das in ANA + 0.1% SDS gelöste Pellet mit Phenol-Chloroform, Chloroform und Äther extrahiert. Nach der abschließenden Ethanol Fällung wurde die RNA in TE gelöst.

Ausbeute : 1-2 µg RNA / entsprechen etwa 2-4% der eingesetzten Gesamtmenge

Für die Markierung wurde die RNA (1µg) lyophilisiert und in 50 mM Trisbase/ pH 9.5 gelöst und 40 min bei 90°C in einer Kapillare hydrolisiert (106). Die danach etwa 50 b langen RNA Fragmente wurden mit 200 µCi γ -³²P ATP in 10 mM MgCl₂/ 5 mM DTT/ 40 mM Tris mit 20 U Polynukleotidkinase für 60 min bei 37°C inkubiert. Das Reaktionsgemisch wurde dann an einer DEAE Sephacell Säule (0.5x4 cm) fraktioniert. Nach dem Auftragen der Probe in 0.3 M NaCl wird mit 0.3 M NaCl ausgiebig gewaschen und die gebundene RNA dann mit 1 M NaCl eluiert. Die RNA läßt sich danach mit Ethanol fällen.

Ausbeute : $2-5 \times 10^7$ cpm/µg RNA

Die Hybridisierung der RNA mit filtergebundener DNA (106) erfolgte in 5xSSC/ 50% Formamid/ 10 mM EDTA/ 0.1% SDS/ 50µg/ml E.coli RNA mit 2.5×10^5 cpm/cm² markierter RNA für 36 Std bei 37°C. Gewaschen wurde danach 2 Std bei 65°C mit

2xSSC, 30 min bei 50°C mit 0.1xSSC und 30 min bei 37°C mit 2xSSC/ 2µg/ml RNase A. Abschließend wurde mehrmals mit 2xSSC bei RT gewaschen, die Filter getrocknet und exponiert.

2.6 Konstruktion der Genbanken

Spaltung und Fraktionierung der genomischen DNA

100-200 µg genomische DNA wurden in 1 ml mit 50-100 U Eco RI in unterschiedlichen Relationen zueinander für 30 min bei 37°C verdaut. Nach dreimaliger Phenolextraktion und Ätherextraktion wurde die DNA Ethanol gefällt und in 1 ml 100 mM NaCl/ 10 mM Tris pH 7.5/ 1mM EDTA gelöst. Je 300 µl (=30-70 µg) wurden 10 min bei 65°C erhitzt und dann auf einen 10-40% linearen Sucrosegradienten im gleichen Puffer (87) aufgeladen. Der Gradient wurde nach der Zentrifugation (SW 27/ 20°C/ 27 krpm/ 24 Std) durch Auspumpen nach oben fraktioniert und die Länge der DNA in den Fraktionen durch Gelelektrophorese bestimmt. Fraktionen mit Fragmenten im Bereich 15-20 kb wurden gepoolt, gegen TE dialysiert und alkoholgefällt.

Isolierung der λ Charon 4A Arme

200 µg λ Ch4A DNA wurde in 1 ml 100 mM Tris pH 7.6/ 25 mM MgCl₂ cyclisiert, auf 50 mM NaCl gebracht und mit 200 U Eco RI 3 Std lang bei 37°C verdaut. Je 60 µg der vollständigen Spaltung wurden auf einem Sucrosegradienten (10-40%) getrennt und fraktioniert. Fraktionen mit Fragmenten von 11, 19 und 31 kb Länge wurden gepoolt, dialysiert und mit Alkohol gefällt.

Ligation der Genbanken

Die Ligationsbedingungen wurden in Anlehnung an (87) gewählt. In einem Gesamtvolumen von 30 μ l 66 mM Tris/ pH7.6/ 6.6 mM $MgCl_2$ / 10 mM DTT/ 0.4 mM ATP und 0.1 U/ μ l Ligase wurden folgende Mengen DNA ligiert :

| Genbank | genom. DNA | λ Arme |
|-----------|-------------|------------------|
| 'Sterna' | 4.2 μ g | 5.1 μ g |
| 'Dediff' | 4.2 μ g | 5.1 μ g |
| Kontrolle | -- | (170 μ g/ml) |

Die Ligationen wurden nach 20 Std bei 15°C alkoholgefällt und in 20 μ l TE sehr vorsichtig gelöst.

In Vitro Packaging

Der In vitro packaging Extrakt wurde nach (92) hergestellt. BHB 2688 und BHB 2690 wurden ausgehend von Einzelkolonien auf LBM Platten dicht ausgestrichen, bei 32°C inkubiert und der Bakterienrasen mit L Medium abgewaschen. Je 300 ml L Medium wurden in 3 l Kolben bis zu einem $OD_{575}=0.07$ angeimpft und bis $OD=0.3$ sehr stark bei 32°C geschüttelt (1.5-2 Std). Ohne Schütteln wurde danach für 15 min bei 45°C die Phagenbildung induziert und danach für 2.5 Std bei 38.5°C nochmals unter Schütteln weiterinkubiert. Die Induktion wurde getestet durch Zugabe von Chloroform zu einer kleinen Menge der Kultur was zu einer Lyse innerhalb etwa 1 min führen sollte. Die Zellen wurden danach abzentrifugiert und in 500 ml M9 Medium aufgenommen. 80 ml Aliquots pro Petrischale (ϕ 15 cm) wurden bei 0°C mit einer UV-Lampe bestrahlt. Die nötige Zeitdauer wurde vorher empirisch ermittelt durch die Bestimmung der Abnahme des Phagentiters nach Bestrahlung. Es wurde eine Zeitdauer gewählt, die zu einer Abnahme des Titers um einen Faktor 10^8 führt (85). Die Zellen wurden danach nochmals zentrifugiert und in 10 ml 40 mM Tris pH8/ 10 mM Spermidin/ 10 mM

Putrescin/ 0.1% Mercaptoethanol/ 7% DMSO suspendiert und nochmals zentrifugiert. Das Zellpellet (etwa 1.2-1.5 ml) wurde mit 0.3 ml des gleichen Puffers verrührt, die Suspension dann auf 5 mM ATP gebracht und auf 2 ml aufgefüllt. 20µl Portionen wurden in flüßigem Stickstoff eingefroren und bei -70°C aufbewahrt.

5µl des konzentrierten Ligationsansatzes wurde mit 20µl Packaging Extrakt verrührt und 45 min bei 37°C inkubiert. Ein weiteres Aliquot des Extrakts wurde langsam bei 0°C getaut und auf 10 mM MgCl₂ und 10 µg/ml DNase I gebracht. Nach dem Verrühren mit dem Packaging Ansatz und Inkubation für weitere 45 min wurde durch Zugabe von 500 µl TMG und 1 Tropfen Chloroform die Reaktion gestoppt. Die Zelltrümmer wurden abzentrifugiert, der Titer des Überstandes bestimmt und die Phagen danach in einem kleinen CsCl Stufengradienten (SW 41) nachgereinigt.

Bei der Konstruktion der Genbanken und deren Screening wurden die Bedingungen der Sicherheitsstufen L3/B2 befolgt (85).

Amplifikation der Genbanken

Zur gleichmäßigen Vermehrung der Sterna Genbank wurden insgesamt 7.6×10^5 pfu (etwa 6.5 Genomequivalente) auf 80 LBM/DT Platten (φ 15 cm) verteilt, was einer Dichte von 10.000 pfu/ Platte entspricht. Die Topagarschicht mit den Plaques wurde abgekratzt und die Platten mit 3-4 ml TMG nachgewaschen. 20 ml Chloroform wurden zu der gesamten Suspension addiert die dann sehr vorsichtig gerührt wurde. Der Agar wurde nach mehreren Stunden abzentrifugiert, der Überstand mit den Phagen abgenommen und mit 10% PEG 6000/ 0.3 M NaCl 20 Std bei 0°C gefällt. Das Phagenpellet wurde mit Chloroform extrahiert und der Überstand getitert.

Analyse des Backgrounds der Genbanken

Die inneren Fragmente des Phagen λ Ch4A können an ihrer genetischen Konstitution erkannt werden: 6 kb = lac⁺, 7 kb = bio⁺. Das lac-Fragment läßt sich durch Plattieren der Phagen auf XGal (40 mg/l) supplementierten LBM/DT Platten an den grünlichblauen Plaques erkennen. Das bio-Fragment wird durch die Ausbildung von Ringen aus schneller wachsenden Bakterien um die einzelnen Plaques erkannt, wenn die Plattierung auf Minimalmedium Platten erfolgt.

2.7 Screening der Genbanken

Bestimmung des Phagentiters der Genbanken

Eine Übernachtskultur von DP50supF in TBMM/DT wurde in 1/10 volumen M-Lösung suspendiert und zu 100 μ l Bakterien versch. Verdünnungen der Phagenlösungen in TMG addiert. Nach 10 min bei 32°C wird 3.5 ml M-Agarose (46°C) zugegeben und auf LBM/DT Platten verteilt. Nach 15 Std lassen sich die Phagenplaques auszählen.

Plattieren der Genbanken und Hybridisierung

200.000 pfu der Genbanken wurden mit 5 ml DP50supF Zellen in M Lösung 10 min inkubiert und in 250 μ l Aliquots mit 7.5 ml M Agarose auf LBM/DT Platten (ϕ 15 cm) in einer Dichte von 10.000 pfu/ Platte verteilt und 15 Std bei 32°C bebrütet. Zum Transfer der Phagen wurden Nitrocellulosefilter auf die Platten aufgelegt und durch Einstiche mit einer sterilen Nadel markiert. Nach 15 min wurden die Filter abgenommen und getrocknet. Duplikate wurden auf die gleiche Weise hergestellt (30 min). Die Phagen und Phagen DNA wird danach denaturiert indem man die Filter auf 200 mM NaOH/ 1.5 mM NaCl für 15 Sekunden mit der DNA Seite nach oben

schwimmen läßt. Auf 400 mM Tris/ pH 7.5/ 2xSSC wird auf die gleiche Weise neutralisiert und auf 2xSSC gewaschen. Nach dem Trocknen der Filter werden diese bei 90°C für 3 Std gebacken. Je 10 Filter wurden mit 2 frischen Nitrocellulose-filtern in einer Tüte mit der nicktranslatierten Probe nach der Sarkosylmethode hybridisiert.

Reinigung positiver Klone

Bereiche auf den Platten die positiven Hybridisierungssignalen entsprachen konnten mit Hilfe der Markierungen wiedergefunden werden. Plaques aus diesen Bereichen wurden einzeln auf Platten mit einem Rasen von DP50supF gepickt. Für den Rasen wurden 100 µl 10-fach konzentrierte Zellen ausplattiert. Nach der Inkubation wurden von diesen Platten wieder Filterreplikas hergestellt und hybridisiert. Positive Plaques wurden dann ausgestochen, in TMG mit einem Tropfen Chloroform einige Zeit bei RT geschüttelt und danach mit einer Impföse die Phagen auf einen Rasen ausgestrichen. Individuelle Plaques wurden wieder auf einen Rasen gepickt, inkubiert und die Filterreplikas erneut hybridisiert. Positive Plaques werden erneut isoliert, ausgestrichen, gepickt und hybridisiert. Die Reinigung wird solange fortgesetzt bis fast 100% der gepickten Plaques positiv reagieren.

Amplifikation des Phagentiters

Um einen höheren Titer für die DNA Präparation zu erhalten wurden pro Platte (Ø 15 cm) 10.000 pfu der isolierten Klone angeimpft, was zu einer konfluenten Lyse führt. 5-10 ml TMG wurden danach auf der Platte verteilt und bei 4°C langsam geschüttelt. Nach Zugabe eines Tropfen Chloroforms kann die Phagenlösung getitert werden und ergibt eine 100-1000 fache Amplifikation des Ausgangstiters.

2.8 Klonierung in Plasmiden

Klonierung in pBR 322

1 μ g pBR 322 wurde mit dem entsprechenden Restriktionsenzym (Eco RI, Bam HI, Hind III) vollständig gespalten. Die zu klonierenden Fragmente wurden mit dem Vektor (Endkonzentration 30-100 μ g/ml) in 10 μ l 1x Ligasepuffer (66 mM Tris/ pH7.6/ 6,6 mM $MgCl_2$ / 10 mM DTT/ 0.4 mM ATP) mit 1-2 U Ligase bei 15°C für 15-30 Std inkubiert. Die Reaktion wurde durch 20mM EDTA gestoppt und direkt für die Transformation von C600 und HB101 verwendet. 100 ml Übernachtskultur wurde 1:100 in frisches Medium verdünnt, bis zu einer $OD_{600}=0.5$ weiterinkubiert und danach zentrifugiert. Das Pellet wurde in 50 mM $CaCl_2$ bei 4°C suspendiert und nach 15 min erneut zentrifugiert. Dieses Pellet wurde in 4 ml 50 mM $CaCl_2$ suspendiert. 0.2 ml der Kompetenten Zellen wurden mit 1-5 μ l des Ligationsansatzes gemischt, 40 min bei 0°C, 2 min bei 37°C und 10 min bei RT inkubiert. Nach der Zugabe von 1 ml L-Medium wurde die Kultur 30 min bei 37°C geschüttelt. 100 μ l der Kultur wurden auf LB platten mit 25 μ g/ml Ampicillin plattiert.

Ausbeute : 0.5-5 x 10⁶ Kolonien/ μ g pBR322

Klonierung in pUC9

pUC 9 (101) wurde ähnlich pBR 322 gespalten, meist mit 0.5 U Alkalischer Phosphatase dephosphoryliert, bei 40-100 μ g/ml ligiert und danach analog pBR322 transformiert. Als Empfängerstamm wurde JM83 verwendet, der zur Selektion auf LB Platten mit 40 mg/ml X-Gal und 200 μ g/ml Ampicillin plattiert wurde. Plasmide mit eingebautem Insert können an der weißen Farbe der Kolonien erkannt werden.

2.9 Klonierung in M13 und Charakterisierung der Klone

Klonierung

Die replikative Form der M13 vektoren mp7, mp8 oder mp9 (65) wurden mit den passenden Restriktionsenzymen gespalten und in einer Konzentration von $2\mu\text{g/ml}$ mit dem zu klonierenden Fragment (etwa 3 facher molarer Überschuß/ $0.01\text{--}0.02\text{ pmol}$) in $70\text{ mM Tris/ pH }7.5/ 0.1\text{ mM ATP}$ für 15 Std bei 15°C ligiert. Transformiert wurde ebenfalls nach der CaCl_2 -Methode, wobei von einzelnen JM103 Kolonien auf Minimalmedium ausgegangen wurde. $100\text{ }\mu\text{l}$ der kompetenten Zellen (in $1/50$ des Ausgangsvolumens der Kultur) wurden mit 2 ng des ligierten M13 Vektors 40 min bei 0°C und 2 min bei 42°C inkubiert. $10\text{ }\mu\text{l}$ 100 mM IPTG , $50\text{ }\mu\text{l}$ 2% X-Gal in DMF und danach 3 ml Softagar (45°C) wurden zugegeben und die Mischung auf LB Platten verteilt. Nach etwa 12 Std lassen sich an der weißen Färbung der Plaques positive Klone erkennen. Zur Kontrolle wurde der M13 Vektor auch ungespalten ($0.5 \times 10^6\text{ Plaques/ }\mu\text{g}$), gespalten ($< 100\text{ Plaques/ }\mu\text{g}$) und religiert ($0.05 - .2 \times 10^6/ \mu\text{g}$) transformiert.

Charakterisierung

Die Phagen in $20\text{ }\mu\text{l}$ des Überstandes einer 2 ml Kultur wurden mit 0.1% SDS lysiert und die DNA nach Zugabe von $1/5$ vol 95% Formamid/ $10\text{ mM EDTA/ }0.1\%$ Bromphenolblau auf einem 1% Agarosegel aufgetrennt. Die Länge der Inserts läßt sich aus dem Vergleich mit M13 Vektor abschätzen.

Um die relative Orientierung von Klonen mit gleichem Insert zu erkennen, wurden je $20\mu\text{l}$ der Kulturüberstände gemischt. 5 min bei 65°C inkubiert, auf RT abgekühlt und auf einem Gel aufgetrennt. Das Auftreten einer zusätzlichen größeren Bande zeigt die Komplementarität der Inserts an. Eine genauere Bestimmung der Insertlänge ist möglich, wenn man je $0.2\text{ }\mu\text{g}$ der 2 Klone mischt, in $10\text{ mM Tris/ pH }7.6/ 50\text{ mM NaCl}$ 10 min bei

65°C inkubiert und langsam auf RT abkühlt. Nach dem Zugabe des gleichen Volumens 2xS1 Puffer (100 mM Naacetat/ pH 4.5/ 4 mM ZnSO₄) wird die einzelsträngige DNA mit 2U S1/μl 20 min bei 37°C abgebaut. Durch den Vergleich mit Marker DNA Fragmenten läßt sich die Länge der verbliebenen doppelsträngigen DNA (= Insertlänge) genau bestimmen.

Konstruktion von M13 Deletionsmutanten

Die Konstruktion der Deletionsmutanten des M13 608A Klonen erfolgte nach (107). 20μg DNA (ds-Form) wurde in 100 μl 50 mM Tris/ pH 7.6/ 1 mM MnCl₂/ 100μg/ml BSA mit 2.5 pg/μl DNase I, die vorher 1 Std im gleichen Puffer bei 0°C inkubiert wurde, 4-16 min bei RT gespalten. Die DNA wurde danach auf einem 1% Low melting point Agarose Gel getrennt, die Bande mit linearisierter DNA ausgeschnitten und daraus extrahiert. Etwa 2 μg der DNA wurde danach mit 25 U Eco RI 6 Std lang gespalten. Der Puffer wurde den Polymerasebedingungen angepaßt und die Enden mit je 40 μM aller dNTP's und 1 U Klenow Fragment in 20 min bei RT aufgefüllt. Mit dem gleichen Volumen 1.6 M NaCl/ 13% PEG 6000 wurde die DNA über Nacht bei 0°C gefällt, mit 60 μl 0.8 M NaCl/ 6.5% PEG gewaschen, mit Phenol und Chloroform extrahiert und danach mit Ethanol gefällt. Die gesamte erhaltene DNA wurde in 10 μl Ligationspuffer gelöst und mit 400 U Ligase 16 Std bei 15°C cyclisiert und in JM 103 transformiert.

Ausbeute : 45 weiße Plaques/ 1-2 μg lineare DNA

2.10 Sequenzierungsmethoden

Methode nach Maxam & Gilbert (64)

Plasmide wurden vor oder nach der Spaltung mit dem entsprechenden Restriktionsenzym durch Chromatographie an Sepharose S-1000 gereinigt und danach die Fragmente (5-50

pmole Enden) mit 0.5U/ μ l Alkalischer Phosphatase in 100 mM Tris/ pH 8.0 30 min bei 37°C dephosphoryliert. Die Reaktion wurde durch 10 mM NTA 20 min bei RT gestoppt und mit 0.3 mM Spermidin 2 min bei 100°C erhitzt. Die DNA Lösung wurde auf

M β -Mercaptoethanol/ 50 mM $MgCl_2$ / 30 mM Tris/ pH 7.7 gebracht und mit 120 μ Ci γ - ^{32}P -ATP (40 pmole) und U Polynukleotidkinase 30 min bei 37°C inkubiert, mit 50 mM EDTA gestoppt, Phenol-Chloroform extrahiert und mit Ethanol gefällt. Die endmarkierten Fragmente wurden nachgespalten, auf einem Acrylamidgel aufgetrennt und die entsprechenden Stücke isoliert. Etwa 30-100.000 cpm endmarkierte DNA wurden auf die 4 Reaktionen verteilt : G, C je 1/6 und G+A, T+C je 1/3 der Gesamtmenge. Reaktionen (G= 4 min/ G+A= 6 min/ T+C= 6 min/ C= 15 min bei 20°C), Fällungen und Piperidinspaltung erfolgten genau nach (64).

Methode nach Sanger (67)

In M13 klonierte Fragmente wurden mit der Dideoxymethode (67) nach einem vereinfachten Reaktionsprotokoll sequenziert. 5-10 ng 'Sequenzierungsprimer' (5'-GTAAACGACG-GCCAGT-3') wurden mit 0.5-1 μ g M13 DNA in 10 μ l 10 mM Tris/ pH 7.5/ 5 mM NaCl 10 min bei 65°C inkubiert und danach langsam auf RT abgekühlt. 40 μ Ci (100 pmole) α - ^{32}P dATP oder α - ^{35}S dATP wurden lyophilisiert und mit dem Annealmix gemischt. 3 μ l basenspezifische Reaktionsmischungen, 2 μ l der Anneal/ dATP Mischung und 0.5-1 U Polymerase/Klenow in 1 μ l wurden durch Zentrifugation gemischt und 15 min bei RT inkubiert, danach 1 μ l Chase Lösung für 15 min zugegeben und mit 3 μ l Stopplösung (95% Formamid/ 10 mM EDTA/ 0.1% Bromphenolblau und Xylencyanol) gestoppt. Bei Verwendung der Schwefelnukleotide wurde 20 min lang inkubiert. Beim Einsatz von 100 pmolen dATP* ergibt sich eine Endkonzentration von 3.3 μ M im Ansatz. Für besonders GC-reiche Sequenzen wurde das Verhältnis dd/d im G und C Mix halbiert (dd/d = 20).

Basenspezifische Reaktionsmischungen (2x Mix) :

| | | G | A | T | C | Chase |
|-------------------|------------|------|-----|------|------|-------|
| ddGTP | (μ M) | 66 | - | - | - | - |
| ddATP | " | - | 300 | - | - | - |
| ddTTP | " | - | - | 117 | - | - |
| ddCTP | " | - | - | - | 66 | - |
| dGTP | " | 1.66 | 33 | 33 | 33 | 2 mM |
| dATP | " | - | - | - | - | 2 " |
| dTTP | " | 33 | 33 | 1.66 | 33 | 2 " |
| dCTP | " | 33 | 33 | 33 | 1.66 | 2 " |
| NaCl | (mM) | 10 | 10 | 10 | 10 | - |
| MgCl ₂ | " | 10 | 10 | 10 | 10 | - |
| Tris/7.5 | " | 10 | 10 | 10 | 10 | 10 |
| DTT | " | 1 | 1 | 1 | 1 | - |
| ddNTP/dNTP | | 40 | 45 | 70 | 40 | |

Sequenzierung des Gegenstrangs : Zur Sequenzierung des Gegenstranges (156) wurde, wie oben beschrieben, der Sequenzierungsprimer an die DNA gebunden und mit je 20 μ M aller 4 NTP's in 10 mM Tris/ pH 7.5/ 5 mM NaCl/ 10 mM MgCl₂/ 1 mM DTT und 1 U Klenow der Komplementärstrang bei RT synthetisiert. Nach 15 min bei RT wurde dabei nochmals 0.5 U Klenow zugegeben und die Reaktion nach weiteren 15 min für 10 min bei 70°C gestoppt. Die DNA wurde in 0.8 M NaCl/ 6.5% PEG 6000 1 Std bei 0°C gefällt, mit der gleichen Lösung gewaschen, Phenol und Äther extrahiert und alkoholgefällt. Die DNA wurde mit 3 ng 'Reverse Sequenzierungsprimer' (5'-CAGGAAACAGCTATGAC-3') in 10 μ l H₂O gemischt und in einer Kapillare 3 min bei 100°C denaturiert. Es wurde in -70°C abgeschreckt, 15 min bei RT inkubiert und nach der Zugabe des Reaktionspuffers wie beschrieben sequenziert.

Sequenzgele

Die Reaktionsprodukte wurden in denaturierenden Acrylamidgelen (45 g/100 ml Harnstoff und Acrylamid/ Bis-

acrylamid = 20:1) aufgetrennt. Die Gellösung wurde vor der Zugabe des Puffers mit Amberlit MB-3 30 min bei RT deionisiert, filtriert und entgast. Es wurden sowohl konventionelle als auch 'sharktooth' Kämme verwendet. Aufgetragen wurden 2-5 μ l (0.5 mm Gele) oder 1-2 μ l (0.3 mm Gele) Probenvolumen, die vorher 5 min bei 95°C denaturiert worden waren. Verwendete Gele :

| | | | |
|---------------|----------------|-----------------|-----------------|
| Größe (LxBxD) | 40x20x0.3 (64) | 80x20x0.3 (109) | 60x20x0.2 (110) |
| Acrylamid | 6-20% | 5-8% | 4-6% |
| Puffer | 1xTBE | 1-1.5xTBE | 1xTBE |
| Temperatur | (50-55°C) | (40-55°C) | 65°C konstant |
| Laufbed. | 2000 V | 2200-2800 V | 2500-3000 V |
| | | 60 W konstant | |

Nach dem Lauf wurden die Gele mit 40 und 80 cm Länge mit 3 MM Papier oder gebrauchten Röntgenfilmen von der Glasplatte abgenommen und bei -70°C exponiert. Die 60 cm langen Gele wurden nach dem Verfahren von Ansorge (110) kovalent an eine Glasplatte gebunden und während des Laufs thermostatisiert (LKB Macromould). Diese Gele wurden nach dem Lauf 10 min in 10% Essigsäure fixiert und 60 min bei 90°C getrocknet.

2.11 Hybrid Selektion und In vitro Translation

Kopplung an DBM-Papier : Für die Hybrid Selektion wurde einzelsträngige DNA von M13 Klonen kovalent an DBM Pulver gebunden. 10 mg ABM-Cellulose (hergestellt nach (76)) wurde mit 3.6 mM NaNO_2 in 1.2 N HCl nach (78) in 30 min bei 4°C aktiviert und mit Wasser und 20 mM Phosphat/ pH 7.4 ausgiebig gewaschen. Etwa 20 μ g M13 DNA (550mp8) wurde in 500 μ l Wasser bei 4°C über Nacht gebunden. Die nicht gebundene DNA wurde durch mehrere Waschungen entfernt : (Je 1 ml) 99% Formamid bei RT; 99% Formamid 5 min bei 65°C; 99% Formamid und mehrmals mit 1xSSC.

Hybrid Selektion : Die an DBM gebundene DNA wurde vor der Hybridisierung 2 Std bei 37°C in 1 ml 50% Formamid/ 5xSSC/ 1x Denhardt/ 1 mg/ml Hefe tRNA/ 1% Glycin/ 0.1% SDS in einem silinierten Eppendorfgefäß durch Rotation gemischt. Nach dem Abzentrifugieren des Pulvers wurde mit versch. RNA Präparationen hybridisiert. 50-100 µg Gesamt RNA wurden in 200 µl 50% Formamid/ 5xSSC/ 1mM EDTA/ 0.1% SDS/ 0.1 mg/ml Hefe tRNA/ 0.1 U/µl RNAsin nach dem Erwärmen auf 70°C für 2 min. 15-30 Std bei 37°C inkubiert. Nicht gebundene RNA wurde abgewaschen und die Überstände verworfen : 3 mal 10 min mit 1 ml 5xSSC/ 2% SDS/ 10 mM EDTA und 3 mal 10 min mit 50% Formamid/ 5xSSC/ 1 mM EDTA/ 0.2% SDS. Eluiert wurde die gebundene RNA in 3 Schritten : 1 mal mit 200 µl 95% Formamid/ 10 mM Tris/ pH 7.5/ 0.5% SDS/ 1 mM EDTA für 30 min bei 37°C (A) und 2 mal mit 200 µl der gleichen Lösung für 2 min bei 70°C (B). Die Überstände A und B wurden alkoholgefällt und in einem kleinen Volumen Wasser gelöst. Methode wurde nach (78) leicht modifiziert.

In Vitro Translation : Isolierte Gesamt RNA (0.1 - 0.5 µg) oder durch Hybrid Selektion gereinigte RNA (20-50% eines Ansatzes) wurde in einem In Vitro Translationssystem aus Kaninchen Reticulocyten nach den Vorschriften der Herstellerfirmen translatiert. Die neu gebildeten Proteine wurden dabei mit ³⁵S-Methionin radioaktiv markiert. durch SDS-Diskelektrophorese (157) aufgetrennt und durch Fluorographie sichtbar gemacht. Für die Kollagenasebehandlung wurde die Hälfte eines Translationsansatzes mit bakt. Kollagenase für 3 min bei 37°C behandelt (80).

2.12 Computerprogramme und Datenbanken

Gelesene Sequenzdaten wurden im Rechner des Rechenzentrums Martinried (VAX 11/782: Digital Equipment) gespeichert und weiterverarbeitet. Für Ambiguitäten der Sequenz wurde folgender Code verwendet :

| | | | |
|---------|---------|-----------|-------------|
| R = A,G | S = A,C | J = C,G,T | N = A,C,G,T |
| Y = C,T | W = G,T | K = A,G,T | M = A,C,G |
| Z = C,G | Q = A,T | L = A,C,T | |

Zur direkten Analyse von DNA Sequenzen wurde das NAQ Programm verwendet, zum Vergleich von Sequenzen untereinander die Programme 'Search', 'Align' und 'Relate' (72). Zu Sequenzvergleichen standen außerdem verschiedene Datenbanken zur Verfügung : EMBL (European Molecular Biology Laboratory, Heidelberg), GenBank (Los Alamos National Laboratory) und NBRF (National Biomedical Research Foundation, Washington). Für Proteindatenvergleiche stand die Dayhoff Datenbank (NBRF) zur Verfügung.

H. LITERATURANGABEN

- (1) Hay, B.
in 'Epithelial-Mesenchymal Interactions', ed. Fleisch-
mayer, R., Billingham, R.E., William & William, Balti-
more (1968)
- (2) Solursh, M.
in 'Cartilage', Vol.2, ed. Hall, B.K.; Academic Press,
New York (1983), 121-141
- (3) Hay, E.D.
in 'Cell biology of the extracellular matrix', ed. Hay,
E.D., Plenum Press, New York, (1981), 379-409
- (4) Kosher, R.A.
in 'Cartilage', Vol 1, ed. Hall, B.K.: Academic Press,
New York (1983), 59-85
- (5) Ramachandran, G.N., Kharta, G.
Nature (1954) 174, 269-270
- (6) Reid, K.B.M.
in 'Collagen in Health and Disease', ed. Weiss, J.B.,
Jayson, M.I.V., Churchill & Livingstone, Edinburgh
(1982)
- (7) Kivirikko, K.I., Myllylä,
in 'Collagen in Health and Disease', ed. Weiss, J.B.,
Jayson, M.I.V., Churchill & Livingstone, Edinburgh
(1982)
- (8) Kühn, K.
in 'Immunochemistry of the Extracellular Matrix', Vol.1,
ed. Furthmayer, H.: CRC Press, Boca Raton (1982) 1-30
- (9) Hessle, H., Engvall, E.
J. Biol. Chem. (1984) 259, 3955-3961
- (10) Mayne, R., Mark von der, K.
in 'Cartilage', Vol.1, ed. Hall, B.K.; Academic Press,
New York (1983) 181-214
- (11) Miller, E.J., Matukas, V.J.
Proc. Natl. Acad. Sci. U.S.A. (1961) 64, 1264-1268
- (12) Butler, W.T., Finch, J.E., Miller, E.J.
J. Biol. Chem. (1977) 252, 639- 643
- (13) Burgeson, R.E., Hollister, D.W.
Biochem. Biophys. Res. Com. (1979) 87, 1124-1131
- (14) Mark von der, K., Menxel van, M., Wiedemann, H.
Eur. J. Biochem. (1984) 138, 629-633

- (15) Reese, C., Mayne, R.
Biochemistry (1981) 20, 5443-5448
- (16) Schmid, T.M., Conrad, H.E.
J. Biol. Chem. (1982) 257, 12444-12457
- (17) Gibson, G.J., Kielty, C.M., Garner, C., Schor, S.L.
Biochem. J. (1983) 211, 417-426
- (18) Hascall, V.C., Hascall, G.K.
in 'Cell Biology of the Extracellular Matrix', ed. Hay,
E.D., Plenum Press, New York (1981) 39-63
- (19) Hascall, V.C., Riolo, R.L.
J. Biol. Chem. (1977) 247, 4529-4538
- (20) Upholt W.B., Vertel, B.M., Dorfman, A.
Proc. Natl. Acad. Sci. U.S.A. (1979) 76, 4847-4851
- (21) Bonnet, F., Legledic, S., Perin, S.P., Jolles, J.,
Jolles, P.
Biochim. Biophys. Acta. (1983) 743, 82-90
- (22) Heinegaard, D., Hascall, V.C.
J. Biol. Chem. (1979) 254, 927-934
- (23) Caplan, A.J.
in '39th-Sympos. of the Soc. Develop. Biol.', ed. Sub-
telny, S., Abbott, U.K.; A.R. Liss, New York (1981) ??
- (24) Kimata, K., Oike, Y., Ito, K., Karasawa, K., Suzuki, S.
Biochem. Biophys. Res. Com. (1978) 85, 1431-1439
- (25) Noro, A., Kimata, K., Oike, Y., Shinomura, T., Maeda,
N., Yano, S., Takahashi, N., Suzuki, S.
J. Biol. Chem. (1983) 258, 9323-9331
- (26) Shinomura, T., Kimata, K., Oike, Y., Noro, A., Hirose,
N., Tanabe, K., Suzuki, S.
J. Biol. Chem. (1983) 258, 9314-9322
- (27) Heinegaard, D., Hascall, V.C.
J. Biol. Chem. (1974) 249, 4250-4256
- (28) Baker, J.R., Caterson, B.
in 'Glycoconjugate Research', Vol. 1, ed. Gregory,
J.D., Jeanloz, R.W., Academic Press, New York (1979)
329-340
- (29) Hewitt, A.T., Varner, H.H., Martin, G.
in 'Immunochemistry of the Extracellular Matrix', Vol.
I, ed. Furthmayer, H., CRC Press Inc.; Boca Raton
(1981) 135-142
- (30) Poole, A.R., Pidoux, J., Reiner A., Choi, H., Rosen-
berg, L.C.

- J. Cell. Biol. (1984) 98, 54-65
- (31) Clarke, I.C.
J. Anat. (1974) 118, 261-280
- (32) Hunziker, E.B., Schenk, R.K.
J. Cell Biol. (1984) 98, 277-282
- (33) Dorfman, A.
in 'Cell Biology of the Extracellular Matrix', ed. Hay, E.D.; Plenum Press, New York (1981) 115-138
- (34) Olsen, B.R.
in 'Cell Biology of the Extracellular Matrix', ed. Hay, E.D.; Plenum Press, New York (1981) 139-178
- (35) Stockwell, R.A.
in 'Cartilage', Vol.1, ed. Hall, B.K.; Academic Press, New York (1983) 253-280
- (36) Hunter, S.J., Caplan, A.I.
in 'Cartilage', Vol.1. ed. Hall, B.K.; (1983) 89-120
- (37) Strom, S.M., Dorfman, A.
Proc. Natl. Acad. Sci. U.S.A. (1976) 73, 1019-1028
- (38) Solursh, M.
in 'Cartilage', Vol.1, ed. Hall, B.K.; Academic Press, New York (1983) 121-142
- (39) Toole, B.P.
in 'Cell Biology of the Extracellular Matrix', ed. Hay, E.D.; Plenum Press, New York (1981) 254-289
- (40) Mark von der, K., Mark von der, H., Gay, S.
Dev. Biol. (1976) 53, 153-170
- (41) Müller, P.K., Lemmen, C., Gay, S., Gauss, V., Kühn, K.
Exp. Cell Res. (1977) 108, 47-55
- (42) Gauss, V.
Doktorarbeit, Universität München (1978)
- (43) Mayne, R., Elrod, B.W., Mayne, P.M., Linsenmayer, T.F.
Exp. Cell Res. (1984) 151, 171-182
- (44) Duchene, M., Sobel, M.E., Müller, P.K.
Exp. Cell Res. (1982) 142, 317-324
- (45) Adams, S.L., Boettiger, D., Focht, R.J., Holtzer, H., Pacivici, M.
Cell (1982) 30, 373-384
- (46) Paglia, L.M., Wiestner, M., Duchene, M., Ouellette, L.A., Horlein, P., Martin, G.R., Müller, P.K.
Biochemistry (1981) 20, 3523-3527

- (47) Pawlowski, P.J., Brierley, G.T., Lukens, L.N.
J. Biol. Chem. (1981) 256, 7695-7698
- (48) Lehrach, H., Frischauf, A.M., Hanahan, D., Wozney, J.,
Fuller, F., Boedtker, H.
Biochemistry (1979) 18, 3146-3152
- (49) Yamamoto, T., Sobel, M.E., Adams, S.L., Avvedimento, V.E.,
Pastan, I., de Crombrughe, B.
J. Biol. Chem. (1980) 255, 2612-2615
- (50) Yamada, Y., Mudryi, M., Sullivan, M., Crombrughe, B.
J. Biol. Chem. (1983) 258, 2758-2761
- (51) Kurkinen, M., Barlow, D.P., Helfman, D.M., Williams, J.G.,
Hogan, B.L.M.
Nucl. Acid Res. (1983) 11, 6199-6209
- (52) Boedtker, H., Fuller, F., Tate, V.,
Int. Rev. Connective Tissue Res. (1983) 10, 1-63
- (53) Yamada, Y., Avvedimento, E., Mudryi, M., Ohkubo, H.,
Vogeli, G., Pastan, I., deCrombrughe, B.
Cell (1980) 22, 887-892
- (54) Lukens, L.N., Frischauf, A.M., Pawlowski, P.J., Brierley, G.T.,
Lehrach, H.
Nucl. Acid Res. (1983) 11, 6021-6040
- (55) Vuorio, E., Sandell, L., Sheffield, V.C., Vuorio, T.,
Dorfman, A., Upholt, W.B.
Nucl. Acid Res. (1982) 10, 1175-1192
- (56) Sandell, L., Yamada, Y., Dorfman, A., Upholt, W.B.
J. Biol. Chem. (1983) 258, 11617-11621
- (57) Strom C.M., Upholt, W.B.
Nucl. Acid Res. (1984) 12, 1025-1037
- (58) Ninomija, Y., Olsen, B.R.
Manuskript in Vorbereitung
- (59) Duchene, M.
Doktorarbeit, Universität München (1984)
- (60) Upholt, W.B., Vertel, B.M., Dorfman, A.
Proc. Natl. Acad. Sci. U.S.A. (1979) 76, 4847-4851
- (61) Upholt, W.B.
Persönliche Mitteilung
- (62) Prockop, D.J., Kivirikko, K.I., Tuderman, L., Guzman, N.A.
New England Journal of Medicine (1979) 301, 13-69
- (63) Fuller, F., Boedtker, H.

- Biochemistry (1981) 20, 996-1006
- (64) Maxam, A.M., Gilbert, W.
Meth. Enzymology (1980) 65, 499-559
- (65) Messing, J.
Meth. Enzymology (1983) 101, 20-77
- (66) Messing, J., Crea, R., Seeburg, P.H.
Nucl. Acid Res. (1981) 9, 309-321
- (67) Sanger, F., Nicklen, S., Coulson, A.R.
Proc. Natl. Acad. Sci. U.S.A. (1977) 74, 5463-5467
- (68) Adams, S.L., Sobel M.E., Howard, B.H., Olden, K., Yamada, K.M., deCrombrugghe, B., Pastan, I.
Proc. Natl. Acad. Sci. U.S.A. (1977) 74, 3399-3403
- (69) Hu, N., Messing, J.
Gene (1982) 17, 271-277
- (70) Thomas, P.S.
Proc. Natl. Acad. Sci. U.S.A. (1980) 77, 5201-5205
- (71) Overbeek, P.A., Merlino, G.T., Peters, N.K., Cohn, V.H., Moore, G.P., Kleinsmith, L.J.
Biochim. Biophys. Acta (1981) 656, 195-205
- (72) George, D.G., Orcutt, B.C., Dayhoff, M.O., Barker, W.C.
in 'NBRF-Reports', National Biomedical Research Foundation (1982).
Georgetown University, Medical Center, Washington
- (73) Carmichael, G.G.
Electrophoresis (1980) 1, 78-82
- (74) Wiestner, M., Fischer, S., Dessau, W., Müller, P.K.
Exp. Cell Res. (1981) 133, 115-125
- (75) Sorbas, Ch.
Doktorarbeit, Universität München (1984)
- (76) Alwine, J.C., Kemp, D.J., Parker, D.A., Reiser, J., Renatz, J., Stark, G.R., Wahl, G.M.
Meth. Enzymol. (1979) 68, 220
- (77) Goldberg, M.L., Lefton, R.P., Stark, G.R., Williams, J.G.
Meth. Enzymol. (1979) 68, 206
- (78) Adams, S.L., Alwine, J., deCrombrugghe, B., Pastan, I.
J. Biol. Chem. (1979) 254, 4935-4938
- (79) Ricciardi, R.P., Miller, J.S., Roberts, B.E.
Proc. Natl. Acad. Sci. U.S.A. (1919) 76, 4927

- (80) Peterkofsky, B.
Meth. Enzymology (1982) 82, 453-471
- (81) Williams, B.G., Blattner, F.R.
in 'Genetic Engineering - Principles and Methods',
Vol.2, ed. Setlow, J.K., Holoender, A.: Plenum Press.
New York (1980)
- (82) Collins, J.
Meth. Enzymology (1979) 68, 309-325
- (83) Blattner, F.R., Williams, B.G., Blechl, A.E., et.al.
Science (1977) 196, 161-169
- (84) Leder, P., Tiemeier, D., Enquist, L.
Science (1977) 196, 175-177
- (85) Richtlinien zum Schutze vor Gefahren durch in vitro
neukombinierte Nukleinsäuren vom 16.7.1980
Bundesministerium für Forschung und Technologie
- (86) Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J.,
O'Connell, C., Quon, D., Sim, G.K., Efstratiadis, A.
Cell (1978) 15, 678-701
- (87) Kretschmer, P.J., Kaufman, R.E., Coon, H.C., Chen,
M.J.Y., Geist, C.E., Nienhuis, A.W.
J. Biol. Chem. (1980) 255, 3204-3211
- (88) Lewin, B.
in 'Eukaryotic Chromosomes': Wiley&Sons, New York
(1983)
- (89) Clarke, L., Carbon, J.
Cell (1976) 9, 91-99
- (90) Dodgson, J.B., Strommer, J., Engel, J.D.
Cell (1979) 17, 879-887
- (91) Tiemeier, D.C., Tilghman, S.M., Leder, P.
Gene (1977) 2, 173-191
- (92) Hohn, B.
Meth. Enzymology (1979) 68, 299-308
- (93) Wet, J.de, Daniels, D.L., Schweder, J.L., Williams,
B.G., Thompson, K.D., Moor, D.D., Blattner, F.R.
J. Virology (1980) 33, 401-410
- (94) Southern, E.M.J.
J. Mol. Biol. (1975) 98, 503-517
- (95) Darnell, J.E.
Nature (1982) 297, 365-371
- (96) Weisbrod, S.

- Nature (1982) 297, 289-295
- (97) Hamada, H., Kakunaga, T.
Nature (1982) 298, 396-398
- (98) Doerfler, W.
Ann. Rev. Biochem. (1983) 52, 93-124
- (99) Ploeg, L.H.T.vander, Flavell, R.A.
Cell (1980) 19, 947-958
- (100) Waalwijk, C., Flavell, R.A.
Nucl. Acid. Res. (1978) 5, 3231-3236
- (101) Vieira, J., Messing, J.
Gene (1982) 19, 259-268
- (102) Böhm, J.
nicht veröffentlichte Daten
- (103) Mount, S.M.
Nucl. Acid Res. (1982) 10, 459-472
- (104) Breathnach, R., Chambon, P.
Ann. Rev. Biochem. (1981) 50, 349-383
- (105) Long, E.O., David, I.B.
Cell (1979) 18, 1185-1196
- (106) Lepesant, J.A., Kejzlarova-Lepesant, J., Garen, A.
Proc. Natl. Acad. Sci. U.S.A. (1978) 75, 5570-5574
- (107) Hong, G.F.
J. Mol. Biol. (1982) 158, 539-549
- (108) Poncz, M., Solowiejczyk, D., Ballantine, M., Schwarz, E., Surrey, S.
Proc. Natl. Acad. Sci. U.S.A. (1982) 79, 4298-4302
- (109) Smith, D.R., Calvo, J.M.
Nucl. Acid Res. (1980) 8, 2263-2269
- (110) Ansorge, W., Barker, R.
J. Biochem. Biophys. Methods (1984) 9, 33-48
- (111) Biggin, M.D., Gibson, T.J., Hong, G.F.
Proc. Natl. Acad. Sci. U.S.A. (1983) 80, 3963-3965
- (112) Fickett, J.W.
Nucl. Acid Res. (1982) 10, 5303-5318
- (113) Robert, L., Moczar, M.
Meth. Enzymol. (1982) 82, 839-853
- (114) Lerner, R.A.
Nature (1982) 299, 592-596

- (115) Hopp, T.P., Woods, K.R.
Proc. Natl. Acad. Sci. U.S.A. (1981) 78, 3824-3828
- (116) Fraga, S.
Can. J. Chem. (1982) 60, 2606-2610
- (117) Binderman, J., Duksin, D., Havell, A., Katzir, E. (Katchalski), Sachs, L.
J. Cell Biol. (1974) 61, 427-439
- (118) Cleary, E.G., Gibson, M.A.
Int. Rev. Connect. Tis. Res. (1983) 10, 97-209
- (119) Oakes, B.W., Bialkower, B.
J. Anat. (1977) 123, 369-387
- (120) Silbermann K.
The Anatomical Record (1983) 206, 373-383
- (121) Gibson, M.A., Grant, M.E., Jackson, D.S.
Connect. Tis. Res. (1982) 10, 145-160
- (122) Gibson, M.A., Cleary, E.G.
Collagen Rel. Res. (1983) 3, 469-488
- (123) Sear, C.H.J., Grant, M.E., Jackson, D.S.
Biochem. J. (1981) 194, 587-598
- (124) Jelinek, W.R., Schmid, C.W.
Ann. Rev. Biochem. (1982) 51, 813-844
- (125) Eden, F.C., Hendrick, J.P., Gottlieb, S.S.
Biochemistry (1978) 17, 813-844
- (126) Sharp, P.A.
Cell (1981) 23, 643-646
- (127) Wahli, W., Dawid, I.B., Wyler, T., Weber, R., Ryffel, G.U.
Cell (1980) 20, 107-117
- (128) Hirano, H., Yamada, Y., Sullivan, M., deCrombrughe, B., Pastan, I., Yamada, K.
Proc. Natl. Acad. Sci. U.S.A. (1983) 80, 46-50
- (129) Kruczek, J., Doerfler, W.
EMBO J. (1982) 1, 409-414
- (130) McKeon, C., Ohkubo, H., Pastan, I., deCrombrughe, B.
Cell (1982) 29, 203-210
- (131) Felsenfeld, G., Nichol, J., Behe, M., McGhee, J., Jackson, D.
in 'CSH Symp. on quant. Biol.' (1983) 47, 577-584
- (132) Davidson, E.H., Britten, R.J.

- Science (1979) 204, 1052-1059
- (133) Igo-Kemenes, T., Hörz, W., Zachau, H.G.
Ann. Rev. Biochem. (1982) 51, 89-121
- (134) Zehner, Z.E., Paterson, B.M.
Proc. Natl. Acad. Sci. U.S.A. (1983) 80, 911-915
- (135) Aho, S., Tate, V., Boedtke, H.
Nucl. Acid Res. (1983) 11, 5443-5450
- (136) Young, R.A., Hagenbüchle, O., Schibler, K.
Cell (1981) 23, 451-458
- (137) Benyajati, C., Spoerel, N., Haymerle, H., Ashburner, M.
Cell (1983) 33, 125-133
- (138) King, C.R., Piatigorsky, D.
Cell (1983) 32, 707-712
- (139) Schwarzbauer, J.E., Tamkun, J.W., Lemischka, J.R., Hy-
nes, R.O.
Cell (1983) 35, 421-428
- (140) Brickell, P.M., Latchman, D.S., Murphy, D., Willison,
K., Rigby, P.W.J.
Nature (1983) 306, 756-760
- (141) Hagenbüchle, O., Tosi, M., Schniebler, U., Boveri, R.,
Wellauer, P.K., Young, R.A.
Nature (1981) 289, 643-646
- (142) Langridge, P., Feix, G.
Cell (1983) 34, 1015-1022
- (143) Faye, G., Leung, D.W., Tatchell, K., Hall, B.O., Smith,
M.
Proc. Natl. Acad. Sci. U.S.A. (1981) 78, 2258-2262
- (144) Malek, L.T., Eschenfeldt, W.H., Munns, T.W., Rhoades,
R.E.
Nucl. Acid Res. (1981) 9, 1657-1668
- (145) Nabeshima, Y., Fuji-Kuriyama, Y., Muramatsu, M., Ogata,
K.
Nature (1984) 308, 333-338
- (146) Aha, S., Tate, V., Boedtke, H.
Nucl. Acid Res. (1983) 11, 5443-5450
- (147) Maniatis, T., Fitch, E.F., Sambrook, J.
in 'Molecular Cloning - A Laboratory Manual' ; Cold
Spring Harbor Laboratory, New York (1982)
- (148) Gauss, V., Müller, P.K.
Biochem. Biophys. Acta (1981) 652, 39-47

- (149) Birnboim, H.C., Doly, J.
Nucl. Acid. Res. (1979) 15, 1515-1523
- (150) Adams, S.L., Sobel, M.E., Howard, B.H., Olden, K.,
Yamada, K.M., deCrombrughe, B., Pastan, I.
Proc. Natl. Acad. Sci. U.S.A. (1977) 74, 3399-3403
- (151) Weislander, L.
Anal. Biochem. (1979) 98, 305-309
- (152) Yang, R.C.A., Lis, J., Wu, R.
Meth. Enzymology (1979) 68, 176-182
- (153) Stellwag, E.J., Dahlberg, A.E.
Nucl. Acid. Res. (1980) 8, 299-317
- (154) Rigby, P.W.Y., Dieckmann, M., Rhodes, C., Berg, P.
J. Mol. Biol. (1977) 113, 237-251
- (155) R  ther, L.L., Koenen, M., Otto, K., M  ller-Hill, B.
Nucl. Acid. Res. (1981) 9, 4087-4098
- (156) Hong, G.F.
Bioscience Reports (1981) 1, 243-252
- (157) Laemmli, U.K.
Nature (1970) 227, 680-685

Danksagung

Mein Dank gilt Herrn Prof. K. Kuhn für die Möglichkeit, die vorliegende Arbeit in der Abteilung für Bindegewebsforschung am Max-Planck Institut für Biochemie anfertigen zu können. Herrn PD Dr. K. Moritz danke ich für die Unterstützung und das Interesse an dieser Arbeit. Herrn Dr. P. K. Müller sei besonders gedankt für die unermüdliche und tatkräftige Unterstützung, aber auch für nützliche Kritik bei der Anfertigung dieser Arbeit. Allen anderen Mitarbeitern der Abteilung danke ich für praktische Hilfen und fruchtbare Diskussionen, besonders Herrn Dr. C. Sorbas und Dr. I. Oberbauer. Mit besonders großer Geduld ertrugen Frau H. Kadner und A. Pöschl alle meine Launen und die permanente Ungeduld bei der Arbeit, wofür ich mich herzlich bedanken möchte.

Die Anfertigung dieser Arbeit wurde jedoch erst ermöglicht durch die dauernde und selbstlose Unterstützung durch meine Familie, meine Eltern und meine Großeltern, denen ich dafür von Herzen danken möchte.

Lebenslauf

Pöschl Ernst

Geburtstag : 24.2.1954

Vater : Pöschl Robert

Mutter : Pöschl Elsa, geb. Huber

- | | |
|-----------|--|
| 1960-1964 | Grundschule in München |
| 1964-1973 | Besuch des Gisela-Gymnasiums in München |
| 1973 | Abitur |
| 1973-1978 | Studium der Biologie an der Ludwig-Maximilians-Universität in München |
| 1978-1979 | Diplomarbeit am Inst. für Physiologische Chemie der Universität München
(Leiter : Prof. H.G. Zachau)
Titel : Klonierung und Sequenzierung von Satelliten-DNA aus Kalbsthymus |
| 1979 | Diplom im Fach Biologie |
| 1979-1981 | Zivildienst in der Gesellschaft für Strahlen- und Umweltforschung, Neuherberg in der Abt. für Zellchemie |
| 1980 | Eheschließung mit Maria Antonia Schell
Geburt der Tochter Agnes Maria |
| 1981-1984 | Promotion am Max-Planck Institut für Biochemie, Martinsried in der Abt. für Bindegewebsforschung (Leiter : Prof. K. Kühn)
Titel : Struktur und Regulation spezifisch regulierter Gene in Bindegeweben des Huhns |

